A COMPACTIFICATION OF THE MODULI SPACE OF TWISTED HOLOMORPHIC MAPS

I. MUNDET I RIERA AND G. TIAN

Abstract. We construct a compactification of the moduli space of twisted holomorphic maps with varying complex structure and bounded energy. For a given compact symplectic manifold X with a compatible complex structure and a Hamiltonian action of S^1 with moment map $\mu : X \to \mathbf{i}\mathbb{R}$, the moduli space which we compactify consists of equivalence classes of tuples (C, P, A, ϕ), where C is a smooth compact complex curve of fixed genus, P is a principal S^1 bundle over C, A is a connection on P and ϕ is a section of $P \times_{S^1} X$ satisfying

$$\overline{\partial}_A \phi = 0, \qquad \iota_v F_A + \mu(\phi) = c,$$

where F_A is the curvature of A, v is the restriction on C of a volume form on the universal curve over $\overline{\mathcal{M}}_g$ and c is a fixed constant. Two tuples (C, P, A, ϕ) and (C', P', A', ϕ') are equivalent if there is a morphism of bundles $\rho : P \to P'$ lifting a biholomorphism $C \to C'$ such that $\rho^* A' = A$ and $\rho^* \phi' = \phi$. The energy of (C, P, A, ϕ) is $\|F_A\|^2_{L^2} + \|d_A\phi\|^2_{L^2} + \|\mu(\phi) - c\|^2_{L^2}$, and the topology of the moduli space is the natural one. We also incorporate marked points in the picture.

There are two sources of non compactness. First, bubbling off phenomena, analogous to the one in Gromov–Witten theory. Second, degeneration of C to nodal curves. In this case, there appears a phenomenon which is not present in Gromov–Witten: near the nodes, the section ϕ may degenerate to a chain of gradient flow lines of $-\mathbf{i}\mu$.

Contents

Date: 31–4–2004.

1. Introduction

1.1. Let C be a compact smooth complex curve with a volume form ν, and let X be a compact symplectic manifold with a compatible almost complex structure. Suppose that X supports a Hamiltonian action of S^1 with moment map $\mu : X \to \mathbf{i}\mathbb{R}$. A twisted holomorphic map (see [M], where the same object was called twisted holomorphic curve) consists of a principal S^1 bundle P over C, a connection A on P and a section ϕ of the associated bundle $P \times_{S^1} X$, satisfying the equations

$$\overline{\partial}_A \phi = 0 \qquad \text{and} \qquad \iota_v F_A + \mu(\phi) = c, \tag{1.1}$$

where F_A denotes the curvature of A, v is a volume form on C, and $c \in \mathbf{i}\mathbb{R}$ is a constant (see Section 2 for the definition of $\overline{\partial}_A \phi$). The second equation is called the vortex equation. Two triples (P, A, ϕ) and (P', A', ϕ') are said to be equivalent if there is an isomorphism of bundles $\rho : P \to P'$ which lifts an automorphism of the curve C and such that $\rho^* A' = A$ and $\rho^* \phi' = \phi$ (note that in [M] we only consider isomorphisms of bundles lifting the identity, i.e., gauge transformations). The set of isomorphism classes of triples (P, A, ϕ) carries a natural topology, and the resulting topological space $\mathcal{M}$ is what we call the moduli space of twisted holomorphic maps over C. There is a notion of energy for triples of the form (P, A, ϕ), called the Yang–Mills–Higgs functional:

$$\mathcal{YMH}_c(P, A, \phi) := \|F_A\|^2_{L^2} + \|d_A \phi\|^2_{L^2} + \|\mu(\phi) - c\|^2_{L^2}. \tag{1.2}$$

Given a number $K > 0$, define $\mathcal{M}(K)$ to be the subset of $\mathcal{M}$ consisting of isomorphism classes of triples (P, A, ϕ) with energy $\leq K$. The space $\mathcal{M}(K)$ is not compact, but the only source of noncompactness is the bubbling off phenomenon well know in Gromov–Witten theory. Thus, it is not a surprise that essentially the same methods as in Gromov–Witten theory (combined with some standard techniques in gauge theory) allow to construct a compactification of $\mathcal{M}(K)$. This was done in [M], and the resulting compactification was used to construct what we call the Hamiltonian Gromov–Witten invariants of X, which are very analogous to Gromov–Witten invariants but depend not only on a symplectic structure but also on a Hamiltonian action of S^1 (both things up to deformation).

1.2. Now, in the general theory of Gromov–Witten invariants one considers maps from curves C to X representing a given homology class, and only the genus of C is fixed: its complex structure varies along the moduli space of curves of the given genus. Looking for a compactification of the resulting moduli space leads to the notion of stable map introduced by Kontsevich in [Ko], which consist of a compact curve C with nodal singularities and with some marked points $\mathbf{x} = (x_1, \dots, x_n)$, and a holomorphic map $f : C \to X$. The only condition on $(C, \mathbf{x}, f)$ is that any rational component of C on

which f restricts to a trivial map has to have at least 3 exceptional (that is, marked or singular) points. A definition of Gromov–Witten invariants using moduli spaces of stable maps was given independently by several authors [FO, LT, S, R] for general compact symplectic manifolds, extending an earlier definition of Gromov–Witten invariants by Ruan and Tian [RT1, RT2] for semipositive symplectic manifolds. One of the most important features of this theory comes from the fact that the moduli space of stable curves fibres in some sense over the Deligne–Mumford moduli space of stable marked curves $\overline{\mathcal{M}}_{g,n}$. This allows to introduce into the play the rich structure in the cohomology of $\overline{\mathcal{M}}_{g,n}$. A beautiful and simple example which shows how powerful this point of view is is the proof of the associativity of the small quantum product in terms of the geometry of $\overline{\mathcal{M}}_{0,4}$. More generally, the gluing formula or composition axiom encodes the structure in Gromov–Witten invariants inherited by the properties of the cohomology of $\overline{\mathcal{M}}_{g,n}$.

In view of this, it seems desirable to generalize the construction given in [M] replacing the fixed curve C by arbitrary marked stable curves, constructing in this way a compact moduli space which fibres over $\overline{\mathcal{M}}_{g,n}$ and obtaining richer invariants. The purpose of this paper is to make a first step towards this aim: we define stable twisted holomorphic maps (c-STHM for short), we give a notion of energy of a c-STHM (the Yang–Mills–Higgs functional), we define a topology on the set of isomorphism classes of c-STHM's, and we prove that the set of isomorphism classes c-STHM's with bounded energy is compact.

1.3. The definition of c-STHM involves several new features which do not appear in the definition of twisted holomorphic maps. The first one is that on nodal curves one has to consider connections whose pullback to the normalization is not necessarily smooth, but only meromorphic. This means that the connection does not extend to the preimages of the nodes, but its curvature extends continuously. To motivate this, consider a sequence of smooth curves C_u converging to a curve C with a nodal singularity, and choose simple closed curves $\gamma_u \subset C_u$ representing the vanishing cycle. If we take bundles P_u and connections on A_u, it may perfectly happen that the holonomy of A_u along γ_u (which is an element of S^1) converges to an element different from the identity as u goes to ∞. If the curvature of A_u is uniformly bounded in compact sets disjoint with γ_u, passing to a subsequence we obtain a limit bundle P and connection A defined on the smooth locus of C, but A does not extend smoothly in the normalization of C because of the nontrivial holonomy. (The reason why we consider connections whose curvature extends continuously will be shortly clarified.)

An important notion related to meromorphic connections is that of limit holonomy: given a meromorphic connection A on the punctured disk $\mathbb{D}^*$ with a pole at the origin, one can define the (limit) holonomy of A around the origin to be the limit of the holonomies around circles centered in the origin (and oriented in a way compatible with the complex structure) as the radius converges to 0. This limit exists because the curvature of A extends continuously to $\mathbb{D}$ (in fact boundedness is enough). The connection A extends to a connection on the whole disk $\mathbb{D}$ (of type C^1) if and only if the holonomy is trivial (see Corollary 3.2).

Let C' be the normalization of C and let y, y' be the preimages of the singular point $z \in C$. The holonomy of the connection A around y is equal to the inverse of the holonomy around y', because both holonomies are obtained as the limits of the holonomies of A_u along γ_u, one using an orientation and the other using the opposite one.

1.4. Suppose now that we have sections ϕ_u of $P_u \times_{S^1} X$ which satisfy equations (1.1) (this involves a choice of volume form in each curve C_u, but for the moment we will ignore this issue — see Section 5.4). Assume that the norm of the covariant derivatives $|d_{A_u}\phi_u|$ is uniformly bounded away from the vanishing cycles. Then the same standard arguments as in [M] allow to obtain, passing to a subsequence and regauging, a limit triple (P, A, ϕ) defined on the smooth locus of C. Furthermore, one can prove easily that $\|d_A\phi\|_{L^2}$ is bounded and that (P, A, ϕ) satisfies equations (1.1).

An important question to understand about (P, A, ϕ) is the asymptotic behavior of ϕ as we approach the singularity of C. If the holonomy of A around the singular point is trivial, then A extends the normalization C' (because its curvature is uniformly bounded), and then Gromov's theorem on removal of singularities (as proved for continuous almost complex structures in [IS]) proves that ϕ extends also to C'. (At this point elliptic bootstrapping using (1.1) proves that in fact both A and ϕ extend smoothly to C'.) To describe what happens when the holonomy is nontrivial, let us consider the real blow up of C at y and y'. This is a real surface $\widetilde{C}$ whose boundary consists of two circles S_y and $S_{y'}$, called exceptional divisors, which can be identified with circles centered at 0 in the tangent spaces T_yC and $T_{y'}C$. The following is a particular case of Corollary 10.2.

Theorem 1.1. *Any triple (P, A, ϕ) defined over the smooth locus of C which satisfies*

$$\overline{\partial}_A\phi = 0 \qquad \text{and} \qquad \|d_A\phi\|_{L^2} < \infty$$

extends to the real blow up $\widetilde{C}$. The restriction of the extension of ϕ to the exceptional divisors is covariantly constant, and takes values in the set of points which are fixed by the action of the limit holonomy of A around y.

Let $F \subset X$ be the set of fixed points. We say that a meromorphic connection has critical holonomy H around a pole if the set of points of X which are fixed by H is bigger than F. The set of critical holonomies forms a finite subset of S^1. Theorem 1.1 implies, in particular, that if the holonomy of A around y' is not critical then ϕ converges somewhere in the fixed point set as we approach z. When the holonomy of A is a root of unity (for example, if the holonomy is critical) Theorem 1.1 follows from Gromov's theorem on removal of singularities, considering local coverings ramified at y and y' so that the pullback connection has trivial holonomy. In fact, the holonomy of A specifies in this case a structure of orbifold near y and y', and the statement of the theorem is equivalent to saying that the section ϕ extends to the normalisation C' as a section of and orbibundle.

Applying Theorem 1.1 to the limit triple (P, A, ϕ) and recalling that the moment map μ is equivariant we deduce that $\mu(\phi)$ extends continuously to the normalization of C. Then the vortex equation implies that F_A also extends continuously, so the limit connection A is indeed meromorphic.

1.5. In case the norms $|d_{A_u}\phi_u|$ are not bounded one can obtain a meaningful limit object by adding bubbles (rational components) to the curve C. When bubbles are attached away from the nodes the picture looks exactly like in Gromov–Witten theory (see [M]). While the first equation in (1.1) is conformally invariant, the second is not, and as we zoom in the curve (which is what we do to construct the limit object in the bubbles) the connection becomes more and more flat. So both P and A are trivial on bubbles away from nodes, and the only nontrivial object on the bubble is the section ϕ, which now can be seen as a holomorphic map from S^2 to the fibre of $P \times_{S^1} X$ over the point in C at which the bubble is attached.

However, if bubbling off occurs near the nodes, new features appear. In the simplest case the curve C is replaced by the union of the normalisation C' and a rational curve C_0, meeting in two nodes. The limit pair (P, A) need not be trivial on C_0: although the connection A will still be flat, it might have poles (nontrivial holonomy) around the nodes. In this case, the resulting bubble is not quite a rational curve.

To give an example we use the cylindrical model $\mathbb{R} \times S^1$ with coordinates t, θ for the rational curve S^2 minus two points. Consider the real function $H := -\mathbf{i}\mu : X \to \mathbb{R}$. For any real number $l \neq 0$, take a gradient flow line $\psi : \mathbb{R} \to X$ at speed l for the function H (so that ψ satisfies the equation $\psi' = l\nabla H(\psi)$). Let $\phi : \mathbb{R} \times S^1 \to X$ be the map $\phi(t, \theta) := \psi(t)$, and let α be the 1-form $\mathbf{i}l d\theta$. Then $d_A := d + \alpha$ gives a covariant derivative in the trivial bundle P over $\mathbb{R} \times S^1$ and we can look at ϕ also as a section of the associated bundle with fibre X. By convention, we assume that the complex structure in $\mathbb{R} \times S^1$ satisfies $\partial/\partial t = \mathbf{i}\partial/\partial\theta$. With these definitions, A is flat and we have $\overline{\partial}_A \phi = 0$. Indeed, the latter equation takes the following form in terms of the cordinates t, θ:

$$\frac{\partial \phi}{\partial t} = I(\phi)\frac{\partial \phi}{\partial \theta} + l\nabla H(\phi), \tag{1.3}$$

which is clearly satisfied. Note also that equation (1.3) is the equation satisfied by connecting orbits in Floer's complex for the Hamiltonian lH (see [F]).

Taken as a connection over S^2 minus two points, A is meromorphic, with holonomy around the poles equal to $\exp(\pm 2\pi \mathbf{i} l)$. Let $x_\pm$ be the limit of $\psi(t)$ as t goes to $\pm\infty$. The contribuition of a bubble C' to the energy of the limit object is given by the squared L^2 norm of $d_A\phi$ on C'. In the example which we have constructed this is

$$\|d_A\phi\|^2_{L^2} = 2\pi l(H(x_+) - H(x_-)), \tag{1.4}$$

so it is finite.

1.6. The factor $H(x_+) - H(x_-)$ in formula (1.4) is uniformly bounded by the constant $\sup_X H - \inf_X H$. Hence, taking l very small, we can obtain bubbles with as small energy as we wish. This makes a big difference with Gromov–Witten theory, where the energy of a rational curve cannot be arbitrarily small. This plays an important role in the compactness theorem in Gromov–Witten theory, where it is used to bound the number of bubbles in a stable map in terms of the energy.

To get a bound on the number of bubbles in our situation (which is crucial if we want to have a reasonable compactness theorem for twisted holomorphic maps of bounded

energy), we need to control the geometry of nontrivial bubbles with little energy. This is done in Theorem 8.4. In fact, the example constructed above gives a hint of what the general situation is: nontrivial bubbles with little energy can be identified with gradient lines of H, which go upward or downward depending on the holonomy of the connection (specifically, the holonomy has to be *near but not quite* critical, and the direction of the gradient line depends on the side of the nearest critical holonomy in which the holonomy of A is). A particular case of this fact was already proved by Floer in Theorem 5 of [F] (recall that the equation $\overline{\partial}_A\phi = 0$ for bubbles is equivalent to equation (1.3) for connecting orbits in Floer's complex), namely, that in which the critical points of H are isolated and the indices of the limit points as $t \to \pm\infty$ differ by one. The argument given by Floer relies on index computations and works only for generic almost complex structure. Our method of proof is different and is based on an analysis of how much a bubble with low energy deviates from being a gradient line (see Theorems 11.1 and 11.3; a similar statement is given in Theorem 1.2 below in this introduction). On the other hand, in our result we require the hamiltonian H to generate an action of S^1, and from this point of view Floer's result is more general than ours.

Using the results in Theorem 8.4 we can associate to any chain of nontrivial bubbles with little energy a chain of gradient lines, each one going from one critical point to another, with as many components as bubbles. Furthermore, since the bubbles are consecutive all the holonomies have to be the same, so the chain of gradient lines is monotone (either always upward or always downward). Finally, since X is compact, the number of components of a monotone chain has to be bounded.

1.7. Let ϕ' be the extension of ϕ to the blow up $\widetilde{C}$ given by Theorem 1.1. Another important question is whether, in the absence of blow up, the restrictions of ϕ' to each exceptional divisor S_y and $S_{y'}$ coincide (up to gauge transformation). The answer turns out to be no: the orbits to which $\phi'(S_y)$ and $\phi'(S_{y'})$ belong need not be the same. Instead, there is a monotone chain of gradient segments going from one to the other. Here by gradient segment we mean the image of a map $\xi : T \to X$ satisfying the equation $\xi' = -\nabla H(\xi)$, where $T \subset \mathbb{R}$ is any closed interval (not necessarily infinite). A chain of gradient segments is a collection of gradient segments $\xi_1(T_1), \dots, \xi_r(T_r)$ such that for any j the segments $\xi_j(T_j)$ and $\xi_{j+1}(T_{j+1})$ meet at a unique critical point $f_j \in F$, and the chain is said to be monotone roughly speaking if $H(f_j)$ decreases with j.

This makes another difference with Gromov–Witten theory, where the analogous question is certainly true. The key point is that the diameter of the image of a holomorphic map $f : C \to X$, where C is any long cylinder, can be bounded in terms of the energy but independently of the length of C. This follows from the exponential decay of the energy density $|df|$ as we go away from the boundary of the cylinder. To understand why there is such exponential decay, we write the equation $\overline{\partial} f = 0$ as an evolution equation $f_t(\tau) = L(\tau)f(\tau)$, where $L(\tau)$ is an elliptic operator on S^1 which is skew symmetric up to compact operators. Then the nonzero spectrum of $L(\tau)$ stays at distance $\geq \sigma > 0$ from 0, so the function $|f_t(\tau)|^2_{L^2(S^1)}$ decays as $e^{-\sigma\tau}$ (note that $|df|^2 = 2|f_t|^2$). This is

something very general (see for example Chapter 3 in [D] for the case of instanton Floer homology).

If we replace the equation $\overline{\partial} f = 0$ by $\overline{\partial}_A \phi = 0$, then the spectra of the operators $L(\tau)$ are shifted by an amount depending on the holonomy of A. And when the holonomy aproaches a critical value, nonzero eigenvalues of $L(\tau)$ can approach arbitrarily 0, so one does not get exponential decay for $|d_A\phi|^2$. The standard way to study the geometry of ϕ in such a situation is to *break* ϕ in two pieces ψ and ϕ_0, in such a way that ψ is spanned by the small eigenvalues of the operator and ϕ_0 is spanned by the big ones. Then ϕ_0 does decay exponentially and so ψ controls the geometry of ϕ away from the boundary of the cylinder. Let us explain in concrete terms how this works in our context.

Consider for simplicity the case in which the holonomy of A is nearly trivial but not quite. Suppose that $C = [-N, N] \times S^1$ and that $P \to C$ has been trivialized in such a way that d_A is approximatedly $d + \lambda\theta$, where $\lambda \in \mathbf{i}\mathbb{R} \setminus \{0\}$ is very small. Suppose that $\phi : C \to X$ has everywhere very little energy $|d_A\phi| < \epsilon$. We can then define $\psi : [-N, N] \to X$ by setting $\psi(t)$ to be the center of mass of $\phi(t, S^1)$, and $\phi_0 : C \to TX$ by the condition $\phi(t, \theta) = \exp_{\psi(t)} \phi_0(t, \theta)$. The following theorem, which combines parts of Theorem 11.1 and Theorem 11.3, gives the main properties of ϕ_0 and ψ when ϕ is a solution of the equations.

Theorem 1.2. *There are constants $K > 0$ and $\sigma > 0$ with the following property. Suppose that $v = f dt \wedge d\theta$ is a volume form on C such that $f(t) < \eta e^{-(N-|t|)}$. Assume that the pair (A, ϕ) satisfies the equations*

$$\overline{\partial}_A \phi = 0 \qquad \text{and} \qquad \iota_v F_A + \mu(\phi) = c.$$

Then we have $|\phi_0(t, \theta)| \leq K e^{-\sigma(N-|t|)}$ and

$$|\psi'(t) + \mathbf{i}\lambda \nabla H(\psi(t))| \leq K e^{-\sigma(N-|t|)} (\epsilon + |\lambda| + \eta)^{1/4}. \tag{1.5}$$

The condition on the volume form is natural, since we will consider cylinders like C as conformal models of curves of the form $\{xy = \delta\} \subset \mathbb{C}^2$ near the origin, with the metric induced by $\mathbb{C}^2$. The exponential bound on $|\phi_0|$ implies that away from the boundary of N the function ϕ can be approximated by ψ, and equation (1.5) tells us that ψ flows near a gradient line at speed $-\mathbf{i}\lambda$.

1.8. To study the behavior of the solutions of (1.1) when C degenerates to a nodal curve, we take cylinders as conformal models, and both ϵ and η go to 0. Hence, if $\lambda \to 0$ as well, (1.5) implies that the map ψ looks more and more like a gradient line. The following result, which is Theorem 12.1, makes this fact precise. The actual meaning of convergence of maps to a chain of gradient segments is given in Section 6.1. A good approximation of this notion is the convergence of subsets in the Hausdorff metric.

Theorem 1.3. *Suppose that $\{\psi_u : T_u \to X, l_u, G_u\}$ is a sequence of triples, where ψ_u is a smooth map, $T_u \subset \mathbb{R}$ is a finite interval, each l_u is a nonzero real number and each $G_u > 0$ is a real number. Suppose that for some $\sigma > 0$ and any u and $t \in T_u$ we have*

$$|\psi_u'(t) - l_u \nabla H(\psi_u(t))| \leq G_u e^{-\sigma d(t, \partial T_u)}. \tag{1.6}$$

Suppose also that $G_u \to 0$ *and that* $l_u \to 0$*. Passing to a subsequence, we can assume that* $l_u|T_u|$ *converges somewhere in* $\mathbb{R} \cup \{\pm\infty\}$ *(here* $|T_u|$ *denotes the length of* T_u*). Then we have the following.*

(1) *If* $\lim l_u|T_u| = 0$ *then* $\lim \operatorname{diam} \psi_u(T_u) = 0$*.*
(2) *If* $\lim l_u|T_u| \neq 0$*, define for big enough* u *and for every* $t \in S_u$ *the rescaled objects* $S_u := l_u T_u$ *and* $f_u(t) := \psi_u(t/l_u)$*. There is a subsequence of* $\{f_u, S_u\}$ *which converges to a chain of gradient segments* $\mathcal{T}$ *in* X*.*

This theorem is almost evident when the critical points of H (that is, the fixed points F) are isolated. In this case it suffices to pay attention to the complementary of a small neighbourhood of F (where, when u is big enough, $l_u\nabla H$ dominates the error term and hence the vector field ψ_u' is almost equal to $l_u\nabla H$), and then shrink the neighbourhood to F. The main difficulty appears when the points of F are not isolated and we want to prove that the limit object is connected: this essentially amounts to proving that the portion of $\psi_u(S_u)$ which lies near F has small diameter, which is not at all obvious (near F the vector field is small, so there is no hope to control pointwise the error term using ∇H — note that we also expect the preimage of a portion of $f_u(S_u)$ near F to become longer and longer as u goes to infinity, so the path f_u may wander slowly near F but for a long time). It is for proving an estimate on the diameter of the intersection of $f_u(S_u)$ with small neighbourhoods of F set (see Lemma 12.3) that we use an exponential bound in the error term.

1.9. In view of all the preceding observations, it is clear that the objects which we should take as limits of solutions of (1.1) over smooth curves degenerating to nodal curves have to be of a mixed nature, combining two dimensional objects (holomorphic sections) with one dimensional objects (gradient flow lines). A very similar thing happens in other related moduli problems. In [CT] the second author and J. Chen construct a compactification of the moduli space of harmonic mappings from compact surfaces to compact Riemannian manifolds, and the limiting objects are a combination of 2-dimensional harmonic maps and geodesics (their one dimensional version). Another instance is the approach suggested by Piunikhin, Salamon and Schwarz [PSS] to proving the equivalence of Floer and quantum cohomology rings using *spiked disks*.

The actual definition of c-STHM's (see Section 5.5) incorporates two additional features. The first one is that we consider marked points in the curves. Following the philosophy of Gromov–Witten theory, we treat marked points and nodes on an equal basis, so in particular we allow poles of the meromorphic connections on marked points. The other feature is some information, for each node of the curve, on how to define the bundle in a smoothening of node (this will be relevant for doing *gluing*, which will be addressed in a future paper). We call this information gluing data (see Section 3.4 for the precise definition). We also give a notion of equivalence between c-STHM's, we define a topology on the set of isomorphism classes of c-STHM's (see Section 6.5), and define the Yang–Mills–Higgs functional $\mathcal{YMH}_c$ for c-STHM's (which is essentially like (1.2)). The main theorem of the paper is then the following (see Theorem 9.1).

Theorem 1.4. *Let g and n be nonnegative integers satisfying $2g+n \geq 3$. Let $K > 0$ be any number, and let $c \in \mathbf{i}\mathbb{R}$. Let $\{\mathcal{C}_u\}$ be a sequence of c-stable twisted holomorphic maps of genus g and with n marked points, satisfying $\mathcal{YMH}_c(\mathcal{C}_u) \leq K$ for each u. Then there is a subsequence $\{[\mathcal{C}_{u_j}]\}$ converging to the isomorphism class of another c-stable twisted holomorphic map $\mathcal{C}$. Furthermore, we have*

$$\lim_{j\to\infty} \mathcal{YMH}_c(\mathcal{C}_{u_j}) = \mathcal{YMH}_c(\mathcal{C}).$$

1.10. **Hamiltonian Gromov–Witten invariants.** The main application of the results in this paper is the construction of Hamiltonian Gromov–Witten invariants coupled to gravity for compact symplectic manifolds, extending the construction given in [M]. This will appear in [MT], and will be based on the compactness result proved here and the technique of virtual moduli cycles. The invariants defined in [M] where obtained by integrating cohomology classes in the moduli space $\mathcal{N}$ of twisted holomorphic maps with a fixed curve C and representing a given class in $H_2^{S^1}(X)$. In [MT] we will construct the moduli space $\overline{\mathcal{N}}_{g,n}$ of c-STHM's of genus g and n marked points and we will define the Hamiltonian Gromov–Witten invariants coupled to gravity. The main new feature compared to [M] (beyond removing the semipositivity conditions imposed on the manifold and the action) is that the moduli space fibres over the Deligne–Mumford moduli space $\overline{\mathcal{M}}_{g,n}$. This implies that the resulting invariants satisfy a gluing axiom similar to the one appearing in Gromov–Witten theory.

1.11. **Contents.** We now briefly summarize the remaining sections of the paper. Sections 2 to 4 are preparatory. In Section 2 we recall the basic objects in the geometry of Hamiltonian almost Kaehler fibrations (covariant derivatives, d-bar operators and minimal coupling form). In Section 3 we define meromorphic connections on bundles over punctured Riemann surfaces, and in Section 4 we recall the definition of orbibundle over orbisurfaces. These sections are included in order to fix notations in a way suitable for our purposes, and there is essentially nothing new in them.

Sections 5 to 9 are devoted to defining c-STHM's and stating and proving the main theorem on compactness. In Section 5 we define c-STHM's and isomorphisms between them, and in Section 6 we define a topology on the set of isomorphism classes of c-STHM's. Section 7 is devoted to the Yang–Mills–Higgs functional, which plays the role of energy for c-STHM's. In Section 8 we prove that a bound on the energy of a c-STHM imposes a bound on the number of irreducible components of the underlying curve. This plays an important role in the compactness theorem, which is the main result of the paper and is stated and proved (modulo technical details) in Section 9.

All technical details of the proof of the compactness theorem are given in Sections 10 to 12. In Section 10 we prove an exponential decay for pairs (A, ϕ) satisfying $\overline{\partial}_A\phi = 0$ and having finite energy on punctured disks. As a Corollary we prove Theorem 1.1 in this introduction. We also prove that the energy of pairs on long cylinders decays exponentially provided the holonomy of the connection stays away from the critical residues. In Section 11 we study pairs on long cylinders with nearly critical holonomy, and we prove that in these pairs the section stays near a curve in X which satisfies up

to normalization and with some error the equation for gradient lines (this is essentially Theorem 1.2). Finally, in Section 12 we prove Theorem 1.3.

1.12. **Notation.** We fix here some notation and conventions which will be used everywhere in the paper. We denote by X a compact symplectic manifold with symplectic form ω. The manifold X supports an effective Hamiltonian action of S^1 with moment map $\mu : X \to (\mathrm{Lie}\, S^1)^*$. We also take a S^1-invariant almost complex structure I on X, compatible with the symplectic form ω in the sense that $g := \omega(\cdot, I\cdot)$ is a Riemannian metric.

We identify $\mathrm{Lie}\, S^1 \simeq \mathbf{i}\mathbb{R}$ with its dual $(\mathbf{i}\mathbb{R})^*$ by using the pairing $\langle a, b\rangle := -ab$ for any $a, b \in \mathbf{i}\mathbb{R}$. Thus in the rest of the paper we will assume that μ takes values in $\mathbf{i}\mathbb{R}$. Let $\mathcal{X}$ be the vector field on X generated by the infinitesimal action of $\mathbf{i} \in \mathbf{i}\mathbb{R}$. Thus the infinitesimal action of any element $\lambda \in \mathbf{i}\mathbb{R}$ gives rise to the vector field $-\mathbf{i}\lambda\mathcal{X}$. We will denote $H := -\mathbf{i}\mu$. Using our convention, we have $H = \langle \mu, \mathbf{i}\rangle$, so that $dH = \langle d\mu, \mathbf{i}\rangle = \iota_{\mathcal{X}}\omega$ by the definition of the moment map. It follows that the gradient ∇H is equal to $I\mathcal{X}$.

Define, for any pair of points $x, y \in X$ the pseudodistance

$$\mathrm{dist}_{S^1}(x, y) := \inf_{\theta \in S^1} d(x, \theta \cdot y),$$

where d denotes the distance in X defined by the Riemannian metric. Also, if $M \subset X$ is any subset, define

$$\mathrm{diam}_{S^1} M := \sup_{x,y \in M} \mathrm{dist}_{S^1}(x, y).$$

Finally, if $A, B \subset X$ are subsets, we define

$$\mathrm{dist}_{S^1}(A, B) := \inf_{x \in A,\ y \in B} \mathrm{dist}_{S^1}(x, y).$$

Whenever we talk about (marked) curves we will implicitly mean that they are compact connected complex curves with at most nodal singularities and eventually with some marked points. These will be a list of smooth labelled points $x_1, \ldots, x_n$, and we will often denote it by a boldface $\mathbf{x}$. When we write $x \in \mathbf{x}$ we will mean that x is one of the points of the list.

Many of the constants appearing in the estimates will be denoted by the same symbol K, and often the value of K will change from line to line.

2. Almost Kaehler geometry of Hamiltonian S^1 fibrations

Let C be a (nonnecessarily compact) complex curve and let P be a principal S^1 bundle over C. Denote by Y the twisted product $P \times_{S^1} X$ and by $\pi : Y \to C$ the natural projection. Let T^{vert} be the vertical tangent bundle $\mathrm{Ker}\, d\pi \subset TY$. Since the almost complex structure I is S^1-invariant, it defines a complex structure of each fibre of T^{vert}, which we still denote by I. We also denote by g the Euclidean form on T^{vert} defined by the S^1-invariant metric g on X. Let I_C be the complex structure on C (so I_C is an endomorphism of the real tangent bundle of C), and let g_C be a conformal metric on C.

Any connection A on P induces a splitting $TY \simeq \pi^*TC \oplus T^{\mathrm{vert}}$, which we use to define $I(A)$ (resp. $g(A)$) as the sum of π^*I_C and I (resp., π^*g_C and g).

Suppose that ϕ is a section of Y. We define the covariant derivative $d_A\phi$ of ϕ with respect to the connection A to be the composition of $d\phi$ with the projection $\pi^v_A : TY \to T^{\mathrm{vert}}$ induced by A. Hence, $d_A\phi$ is a one form on C with values in the pullback ϕ^*T^{vert}.

Lemma 2.1. *Let S be the interval $[0,1]$, let $\gamma : S \to C$ be a smooth map, and let ϕ be a section of Y. Then $\mathrm{dist}_{S^1}(\phi\gamma(1), \phi\gamma(0))$ is at most the integral over S of $|\langle d_A\phi(\gamma), \gamma'\rangle|$.*

Proof. Take a trivialization $\gamma^*Y \simeq [0,1] \times X$, and consider a gauge transformation $g : [0,1] \to S^1$ which sends γ^*A to the trivial connection. Then we have $\mathrm{dist}_{S^1}(\phi\gamma(1), \phi\gamma(0)) = \mathrm{dist}_{S^1}(g\phi\gamma(1), g\phi\gamma(0))$, and the latter can easily be estimated in terms of the integral of $|\langle d_A\phi(\gamma), \gamma'\rangle|$ (note that by covariance we have $\langle d_A\phi(\gamma), \gamma'\rangle = \langle d_{g^*\gamma^*A}(g\phi), \partial/\partial t\rangle$). $\square$

We can split the space of forms $\Omega^1(C, \phi^*T^{\mathrm{vert}})$ as the sum of the space of holomorphic forms $\Omega^{1,0}(C, \phi^*T^{\mathrm{vert}})$ plus the space of antiholomorphic forms $\Omega^{0,1}(C, \phi^*T^{\mathrm{vert}})$. Let $\overline{\partial}_{I,A}\phi$ be the projection of $d_A\phi$ to $\Omega^{0,1}(C, \phi^*T^{\mathrm{vert}})$. In concrete terms,

$$\overline{\partial}_{I,A}\phi := \frac{1}{2}(d_A\phi + I \circ d_A \circ I_C). \tag{2.7}$$

(When the complex structure on X will be clear from the context, we will simply write $\overline{\partial}_A\phi$.) Let us denote by $\Phi : C \to Y$ the map defined by the section ϕ. It is straightforward to check that

$$\overline{\partial}_{I,A}\phi = 0 \qquad \Longleftrightarrow \qquad \overline{\partial}_{I(A)}\Phi = 0. \tag{2.8}$$

On the other hand, for any section Φ of Y we have

$$|d\Phi|^2_{g(A)} = |d\,\mathrm{Id}_C|^2 + |d_A\phi|^2 = 1 + |d_A\phi|^2. \tag{2.9}$$

Given a connection A on P, there is a canonical way to pick a closed 2-form $\omega(A)$ on Y which restricts to ω on each fiber. This is the **minimal coupling form** (see [GLS]). To give a local description of it we can assume that there is a trivialisation $P \simeq C \times S^1$. Let $\alpha \in \Omega^1(C, \mathbf{i}\mathbb{R})$ be the 1-form corresponding to A with respect to this trivialisation. Let also $\pi_X : Y \to X$ the projection induced by the trivialisation. Then we have

$$\omega(A) = \pi_X^*\omega - d(\pi^*\alpha \wedge \pi_X^*\mu) = \pi_X^*\omega - \pi^*\alpha \wedge \pi_X^*\iota_{\mathcal{X}}\omega - \pi^*F_A \wedge \pi_X^*\mu, \tag{2.10}$$

where F_A is the curvature of A. If A is flat, then $\omega(A)$ coincides with the 2 form $g(A)(\cdot, -I(A)\cdot)$ when restricted to vertical tangent vectors.

Remark 2.2. *The cohomology class represented by $\omega(A)$ is independent of A, and can in fact be identified to the pullback of the class $[\omega - \mu t]$ in equivariant cohomology (this denotes the class represented by the element $\omega - \mu t$ of the Cartan–Weil complex $\mathbb{R}[t] \otimes \Omega(X)^{S^1}$, see for example [GS]). More precisely, if $c : C \to BS^1$ is the classifying map for P, then we can identify $Y \simeq c^*X_{S^1}$ and $[\omega(A)]$ is equal to $c^*[\omega - \mu t]$.*

3. Meromorphic connections on marked nodal surfaces

3.1. **Meromorphic connections on marked smooth curves.** Let $(C, \mathbf{x})$ be a marked smooth complex curve, and let P be a principal S^1 bundle over $C \setminus \mathbf{x}$. For any $x \in \mathbf{x}$ we denote by $T(P, x)$ the set of trivialisations up to homotopy of the restriction of P to a small loop γ around x (since for two different loops γ and γ' the sets of trivialisations of the corresponding restrictions can be canonically identified, $T(P, x)$ is independent of the chosen loop). Orient γ counterclockwise with respect to the natural orientation of C as a Riemannn surface, $T(P, x)$ gets a natural structure of $\mathbb{Z}$-torsor[1], given by the action of gauge transformations defined over γ: if $t : P|_\gamma \to \gamma \times S^1$ is a trivialisation and $\tau := [t]$ denotes its class in $T(P, x)$, then for any $n \in \mathbb{Z}$ we define $n \cdot \tau$ to be the class of the trivialisation $t \circ g$, where $g : P|_\gamma \to P|_\gamma$ is the gauge transformation given by any map $\gamma \to S^1$ of index n.

Any $\tau \in T(P, x)$ defines an extension of P to x: if $D \subset C$ is a disk containing x and having γ as boundary, then we glue P to the trivial bundle over D using any extension of a representative of τ to $D \setminus \{x\}$ as patching function. This extension is well defined up to homotopy. In general, if we pick a collection of local trivialisations

$$\tau \in T(P, \mathbf{x}) := \prod_{x \in \mathbf{x}} T(P, x)$$

then we get an extension of P to C, which we denote by P^τ. We define the **degree** of the bundle P over $(C, \mathbf{x})$ to be the map

$$\deg P : T(P, \mathbf{x}) \to \mathbb{Z}$$

which sends any $\tau \in T(P, \mathbf{x})$ to $\deg P(\tau) := \deg P^\tau$.

We say that a connection A on P is **meromorphic** with poles in $\mathbf{x}$ if its curvature F_A extends to the whole C as a continuous 2-form.

Suppose that A is meromorphic. Chose any metric on C, and fix some $x \in \mathbf{x}$. Let γ_ϵ denote the boundary of a geodesic disk centered on x and of radius ϵ, oriented counterclockwise. Denote by $\operatorname{Hol}(A, \gamma_\epsilon) \in S^1$ the holonomy of A around γ_ϵ. We claim that the limit

$$\operatorname{Hol}(A, x) := \lim_{\epsilon \to 0} \operatorname{Hol}(A, \gamma_\epsilon) \tag{3.11}$$

exists. Indeed, if $\epsilon > \epsilon' > 0$ then the quotient $\operatorname{Hol}(A, \gamma_\epsilon)/\operatorname{Hol}(A, \gamma_{\epsilon'})$ is equal, by Stokes' theorem, to the exponential of the integral of the curvature F_A over the region enclosed between γ_ϵ and $\gamma_{\epsilon'}$. But since F_A extends continuously to x, this integral tends to 0 as $\epsilon \to 0$, so the quotient of the holonomies tends to 1. The same argument proves that the definition of $\operatorname{Hol}(A, x)$ is independent of the chosen metric.

Lemma 3.1. *Let (r, θ) be polar coordinates defined on a small disk D centered on $x \in \mathbf{x}$. Any $\tau \in T(P, x)$ has a representative which extends to give a trivialisation of P on $D \setminus \{x\}$ with respect to which the covariant derivative associated to A can be written as*

$$d_A = d + \alpha + \lambda d\theta, \tag{3.12}$$

[1]Recall that if Γ is a group, a Γ-torsor is a set T with a free left action of Γ.

where $\alpha \in \Gamma(T^*D \otimes \mathbf{i}\mathbb{R})$ *is a* 1*-form of type* C^1*, and* $\lambda \in \mathbf{i}\mathbb{R}$ *satisfes* $\mathrm{Hol}(A,x) = e^{2\pi\lambda}$*. Furthermore, any two choices of trivialisations associated to the same* $\tau \in T(P,x)$ *are related by a gauge transformation which extends to* D*.*

Proof. Take a trivialisation of P on $D \setminus \{x\}$ extending any representative of τ. With respect to this trivialisation we can write $d_A = d + \alpha_0$, so that $F_A = d\alpha_0$. Let

$$\lambda := \lim_{\epsilon \to 0} \frac{1}{2\pi} \int_{\theta \in S^1} \alpha_0(\epsilon, \theta).$$

That this limit exists follows as in the definition of $\mathrm{Hol}(A,x)$ from Stokes' theorem, writing the difference of the integrals on the right hand side for two choices of ϵ as the integral of $d\alpha$ on the corresponding annulus, and then using the fact that $d\alpha_0$ extends continuously to D. By Poincaré's lemma, there exists $\alpha \in \Omega^1(D, \mathbf{i}\mathbb{R})$ of type C^1 such that $d\alpha_0 = d\alpha$. Hence $\beta := \alpha_0 - \alpha - \lambda d\theta$ is closed. But it is also exact, since for any $\epsilon > 0$ we have

$$\int_{\theta \in S^1} \beta(\epsilon, \theta) = \lim_{\epsilon \to 0} \int_{\theta \in S^1} \beta(\epsilon, \theta) = \lim_{\epsilon \to 0} \int_{\theta \in S^1} \alpha_0 - \lambda d\theta = 0$$

(in the first equation we use that β is closed, in the second one that α_0 extends to D and in the third one the definition of λ). Consequently, we can write $\beta = dg$ for some $g : D \setminus \{x\} \to S^1$. Now, the gauge transformation $G := \exp g$ transforms $d + \alpha_0$ into $d + \alpha + \lambda d\theta$. If another gauge transformation G' with winding number 0 transform $d + \alpha_0$ into $d + \alpha' + \lambda d\theta$ and α' extends to D, then we can write $G' = dg'$ (because the winding number of G' is 0) and $dg - dg' = \alpha - \alpha'$. The right hand side is a closed 1-form which extends to x_j, hence $g - g'$ extends to D. □

Corollary 3.2. *Let* P *be a bundle over* $C \setminus \mathbf{x}$ *and let* A *be a meromorphic connection on* P*. Let* $\mathbf{x}' \subset \mathbf{x}$ *be the marked points around which the holonomy of* A *is nontrivial. Then there is a bundle* $P' \to C \setminus \mathbf{x}'$ *with a smooth connection* A' *such that* $\iota^*(P', A') \simeq (P, A)$*, where* $\iota : C \setminus \mathbf{x} \to C \setminus \mathbf{x}'$ *denotes the inclusion.*

Remark 3.3. *In fact, it is even possible to trivialize* P *around* x *in such a way that* d_A *is in* **radial gauge***, that is, it takes the form* $d + \alpha + \lambda d\theta$ *and* $\alpha = \alpha_\theta d\theta$ *is of type* C^1 *(in particular,* α_θ *vanishes at* 0*). Indeed, given a covariant derivative* $d + \alpha + \lambda d\theta$ *as in Lemma 3.1 on the trivial bundle over the punctured disk, we can define* $g : D \to S^1$ *to be the map whose value at* $z \in D$ *gives the parallel transport with respect to* $d + \alpha$ *from the fibre over* z *to that over* 0*. Then* $d + \alpha - g^{-1}dg$ *is in radial gauge. Restricting to the punctured disk, the gauge transformation* g *sends* $d + \alpha + \lambda d\theta$ *to a connection in radial gauge.*

We call $\lambda =: \mathrm{Res}(A, x, \tau) := \mathrm{Res}(A, x, \tau_x)$ the **residue** of A on x with respect to τ. It is straightforward to check that for any $n \in \mathbb{Z}$ we have

$$\mathrm{Res}(A, x, n \cdot \tau) = \mathbf{i}n + \mathrm{Res}(A, x, \tau). \tag{3.13}$$

Remark 3.4. *One can check easily that if* A *is a connection whose curvature* F_A *is uniformly bounded (but not necessarily extends continuously to* C*) then Lemma 3.1 is still true, with the difference that the* 1*-form* α *is in general only continuous. Consequently,*

one can also define in this situation the residue $\mathrm{Res}(A,x,\tau)$ *and formula (3.13) still holds.*

3.2. **Meromorphic connections on marked nodal curves.** Now suppose that $(C,\mathbf{x})$ is a marked curve with nodal singularities. Denote by $\mathbf{z}$ the set of nodes of C. Let $\pi : C' \to C$ be the normalisation map. We denote the preimage by π of the marked points $\mathbf{x}$ with the same symbol $\mathbf{x}$.

If P is a bundle over $C \setminus (\mathbf{x}\cup\mathbf{z})$, by a **meromorphic connection** on P we mean a meromorphic connection on π^*P such that for every node $z \in \mathbf{z}$, denoting by y, y' the preimages of z in the normalization, we have

$$\mathrm{Hol}(A,y) = \mathrm{Hol}(A,y')^{-1}.$$

Define[2] for any such node $T(P,z) := T(\pi^*P,y)\times_{\mathbb{Z}} T(\pi^*P,y')$. A **marked** S^1 **principal bundle** over $(C,\mathbf{x})$ is a pair (P,σ), where P is a principal S^1 bundle over $C\setminus(\mathbf{x}\cup\mathbf{z})$ and

$$\sigma \in T(P,\mathbf{z}) := \prod_{z\in\mathbf{z}} T(P,z).$$

Let C^s be a smoothening of C (hence, C^s is isomorphic to C away from a collection of neighbourhoods of the nodes of C, and near the nodes C^s takes the form $\{xy=\epsilon\}\subset\mathbb{C}^2$ instead of $\{xy=0\}\subset\mathbb{C}^2$). The elements in σ give a unique way (up to homotopy) of extending P to a bundle P^s over $C^s\setminus\mathbf{x}$ (here we identify $\mathbf{x}$ with points in C^s). Then we define the **degree** of (P,σ) to be the map

$$\deg P^\sigma := T(P,\mathbf{x}) \to \mathbb{Z}$$

given by $\deg P^\sigma := \deg P^s$.

A **meromorphic connection** on (P,σ) is a meromorphic connection A on π^*P with poles in $\mathbf{x}\cup\mathbf{y}$ which satisfies the following compatibility condition on each node z: if y,y' are the preimages of z in C' and we take a representative $(\xi,\xi')\in T(\pi^*P,y)\times T(\pi^*P,y')$ of σ_z then

$$\mathrm{Res}(A,y,\xi)+\mathrm{Res}(A,y',\xi') = 0. \tag{3.14}$$

In other words, the sum of the residues in both preimages of the node has to vanish. That this condition is well defined follows from (3.13).

It is straightforward to check that given a connection A on P there is a unique choice of σ with respect to which A is meromorphic (that is, formula (3.14) holds). Using such σ, we define

$$\deg P^A := \deg P^\sigma.$$

[2]If T and T' are two Γ-torsors and Γ is abelian, we denote $T\times_\Gamma T' := T\times T'/\sim$, where the equivalence relation $\sim$ identifies, for any $(t,t')\in T\times T'$ and $g\in\Gamma$, $(g\cdot t,t')\sim(t,g\cdot t')$.

3.3. **Chern–Weil formula.** The following lemma gives a generalisation of the simplest Chern–Weil formula to the case of meromorphic connections on nodal curves.

Lemma 3.5. *Denote by F_A the curvature of a meromorphic connection A on P. For any $\tau \in T(P, \mathbf{x})$ we have*

$$\deg P^A(\tau) = \frac{\mathbf{i}}{2\pi} \int_C F_A + \mathbf{i} \sum_{x \in \mathbf{x}} \operatorname{Res}(A, x, \tau). \tag{3.15}$$

Proof. It is enough to consider the case of smooth C (in the nodal case, pulling back to the normalisation we are led to the smooth case). Hence we assume that C is smooth and that A is a meromorphic connection on a bundle P over $C \setminus \mathbf{x}$. Let $\mathbf{x} = (x_1, \ldots, x_n)$. Pick some $\tau \in T(P, \mathbf{x})$ restriction of P to a small loop around each marked point and define $\lambda_j := \operatorname{Res}(A, x_j, \tau)$. Take disjoint disks $D_1, \ldots, D_n$ centered at the marked points, small enough so that we can apply Lemma 3.1 to get trivialisations of P near each x_j for which the covariant derivative of A takes the form $d + \alpha_j + \lambda_j d\theta$ (here (r, θ) are polar coordinates in D_j). For any j chose a smooth 1-form η_j on D_j which coincides with $d\theta$ away from a small disk $\{r < \epsilon\} \subset D_j$. Now, modify A near each x_j by replacing $d + \alpha_j + \lambda_j d\theta$ by $d + \alpha_j + \lambda_j \eta_j$. In this way we get a smooth connection A' on the bundle P^τ over C. Then using standard Chern–Weil and Stokes we compute

$$\begin{aligned}\deg P(\tau) = \deg P^\tau = \frac{\mathbf{i}}{2\pi} \int_C F_{A'} &= \frac{\mathbf{i}}{2\pi} \int_C F_A + \frac{\mathbf{i}}{2\pi} \sum \lambda_j \int_{D_j} d\eta_j \\ &= \frac{\mathbf{i}}{2\pi} \int_C F_A + \frac{\mathbf{i}}{2\pi} \sum \lambda_j \int_{\partial D_j} \eta_j = \frac{\mathbf{i}}{2\pi} \int_C F_A + \mathbf{i} \sum \lambda_j.\end{aligned}$$

□

3.4. **Gluing data.** Suppose that $(C, \mathbf{x})$ is a smooth curve, P is a principal S^1 bundle over $C \setminus \mathbf{x}$ and A is a meromorphic connection on P. For every $x \in \mathbf{x}$, let S_x be the quotient of $T_x C^* / \mathbb{R}$, where $T_x C^* \subset T_x C$ denotes the nonzero elements and $\mathbb{R}$ acts on $T_x C^*$ by multiplication. The conformal structure on C induces an orientation and a metric on S_x up to constant scalar. Imposing the volume of C to be 2π, we get a well defined metric on S_x. A useful point of view is to look at S_x as the exceptional divisor of the real blowup of C at x.

The pair (P, A) induces a **limiting pair** (P_x, A_x), where P_x is a bundle over S_x and A_x is a connection on P_x in the following way. Pick any conformal metric on C and take a small $\epsilon > 0$. Let $D_\epsilon := \exp_x\{v \mid v \in T_x C^*,\ |v| < \epsilon\}$. Let $D_\epsilon / \mathbb{R}$ denote the quotient by the equivalence relation which identifies y, z if and only if $y = \exp_x u$, $z = \exp_x v$ and $u \in \mathbb{R} v$. We can lift this equivalence to the restriction of P on D_ϵ by using parallel transport (with respect to A) along lines of the form $\{\exp_x \lambda u \mid \lambda \in \mathbb{R}\}$, and we denote the quotient by $P_x := P|_{D_\epsilon} / \mathbb{R}$. We look at P_x as a bundle over S_x using the obvious identification $S_x \simeq D_\epsilon / \mathbb{R}$. To define the limiting connection we proceed as follows: for any $\delta > 0$ smaller than ϵ, let S_δ be the exponential of the circle of radius ϵ in $T_x C$. The composition of inclusion and quotient: $S_\delta \hookrightarrow D_\epsilon \to D_\epsilon / \mathbb{R} \simeq S_x$ is an isomorphism,

and we denote by g_δ its inverse. Using parallel transport g_δ lifts to an isomorphism $P_x \simeq g_\delta^* P|_{S_\delta}$. One checks, using the fact that A is meromorphic, that the pullback connections $g_\epsilon^* A$ converge to a limit connection, which we denote by A_x. (For example, use a trivialisation around x which puts d_A in radial gauge and as $d_A = d + \alpha + \lambda d\theta$, where α extends to x — see Remark 3.3.) The resulting pair (P_x, A_x) is independent of the chosen metric on C. Of course, the holonomy of A_x around S_x coincides with the holonomy $\mathrm{Hol}(A, x)$ of A around x as defined by (3.11).

In other words, the pullback of (P, A) to the real blowup of C at x extends to the exceptional divisor S_x, and the restriction to S_x of the extension is isomorphic to (P_x, A_x).

Now suppose that $(C, \mathbf{x})$ is a nodal curve. Let $z \in C$ be a node and let y, y' be its preimages in the normalisation C' of C. Define the set of **gluing angles** at z to be

$$\Gamma_z := (T_y C' \otimes T_{y'} C')^* / \mathbb{R}.$$

Recall that the nonzero elements of $T_y C' \otimes T_{y'} C'$ specify deformations of C which smoothen the singularity at z (see Section 6.3). On the other hand, the set Γ_z can be identified with the set of isometries $\gamma : S_y \to S_{y'}$ which reverse the orientations.

Let P be a principal S^1 bundle over $C \setminus (\mathbf{x} \cup \mathbf{z})$, where $\mathbf{z}$ denotes the set of nodes. We define the set of **gluing data for** P **at** z to be the set $\Gamma(P, z)$ of pairs (γ, ρ), where $\gamma \in \Gamma_z$ is a gluing angle at z and $\rho : \gamma^* P_{y'} \to P_y$ is an isomorphism of bundles satisfying $\rho^* A_{y'} = A_y$. Clearly, since γ reverses the orientation the condition for ρ to exist is that the holonomy of $A_{y'}$ is the inverse of that of $A_{y'}$. Note that the projection $\pi_z : \Gamma(P, z) \to \Gamma_z$ has a natural structure of principal S^1 bundle. Finally, we define the set of **gluing data for** P to be the product

$$\Gamma(P) := \prod_{z \in \mathbf{z}} \Gamma(P, z)$$

of sets of gluing data at each of the nodes of C. The group of gauge transformations of P acts in an obvious way on $\Gamma(P)$ preserving the gluing angles (in other words, for every node z the induced action on $\Gamma(P, z)$ preserves the map π_z).

Any diagram of the form

$$\begin{tikzcd} P \arrow[r, "g"] \arrow[d] & P' \arrow[d] \\ C \arrow[r, "f"] & C', \end{tikzcd}$$

where g is an isomorphism of principal S^1 bundles and f is a biholomorphism satisfying $f(\mathbf{x}) = \mathbf{x}'$ induces a **morphism of gluing data** $g^* : \Gamma(P') \to \Gamma(P)$ (note that the vertical arrows in the diagrams are not surjections, since the bundles are defined over the set of smooth points of C and C' which are not marked).

A choice of gluing data for P should be understood as a device to specify deformations of P lifting certain smoothings of C. More precisely, suppose that $G \in \Gamma(P)$ consists of gluing data $(\gamma_1, \rho_1), \dots, (\gamma_k, \rho_k)$. The set of gluing angles $(\rho_1, \dots, \rho_k)$ give a real subspace T in the tangent space of $\overline{\mathcal{M}}_{g,n}$ at the point represented by $(C, \mathbf{x})$ (suppose for simplicity

that we only smoothen the singularities and do not modify the complex structures of the irreducible components of C). Then, for each deformation C_t parametrized by some small segment $t \in [0, \epsilon)$ tangent to T, G gives a way to extend (P, A) along C_t, in a *infinitesimally unique* way.

This may seem a little unsatisfactory, since it would be preferable to specify some data which tells how to deform (P, A) along any deformation of $(C, \mathbf{x})$ in $\overline{\mathcal{M}}_{g,n}$. Such a thing is certainly possible if A has trivial holonomy around the preimages of z. In this case, the bundle π^*P (recall that π is the normalisation map) extends over y and y' (essentially by Lemma 3.1) to give a bundle P', and picking an identification of the fibres P'_y and $P'_{y'}$ (which should not be confused with P_y and $P_{y'}$) gives a way to extend P to any smoothing of C at z. Furthermore, the set of possible identifications between P'_y and $P'_{y'}$ is a torsor over S^1.

In view of this we could wonder whether it is possible to chose, when A has nontrivial holonomy H around y and y', a family parametrized by S^1 of sections of the bundle $\Gamma(P, z) \to \Gamma_z$, varying continuously with H (the point is that such a section would give a way to deform P for each direction of smoothings of C). But this is unfortunately not possible, as we now explain. Denote by $\mathcal{H}$, $\mathcal{H}'$ the moduli spaces of connections on the trivial S^1 bundle over S_y, $S_{y'}$. We then have $\mathcal{H} \simeq S^1 \simeq \mathcal{H}'$ canonically, the isomorphism being given by the holonomy. There are universal Poincaré bundles $\mathcal{P} \to S_y \times \mathcal{H}$ and $\mathcal{P}' \to S_{y'} \times \mathcal{H}'$ (with a universal connection A whose restriction to the fibre $S_y \times \{H\}$ has holonomy $H \in S^1$, and similarly there is a connection A' on $\mathcal{P}'$). Consider the bundle $\mathbb{I} := \mathrm{Isom}(\mathcal{P}, \mathcal{P}')$ over $\Gamma_z \times S^1$ whose fibre on (γ, H) is the set of isomorphisms between $\mathcal{P}|_{S_y \times \{H\}}$ and $\gamma_I^* \mathcal{P}'|_{S_{y'} \times \{H^{-1}\}}$, where $\gamma_I(\alpha, H) := (\gamma(\alpha), H^{-1})$, which preserve the connections. What we are looking for is a family of sections $\Sigma_H \subset \Gamma(\mathbb{I}|_{\Gamma_z \times \{H\}})$ for every H and depending continuously on H, at least homeomorphic to S^1. If such a thing existed, it would form a fibration Σ over S^1 with connected fibres, and hence it would admit sections. Any section of Σ would induce a section of $\mathbb{I}$, which is impossible, because the bundle $\mathbb{I}$ is nontrivial (to see that $\mathbb{I}$ is nontrivial observe that the connections A and A' induce a connection $\mathbb{A}$ on $\mathbb{I}$ whose holonomy around $\Gamma_z \times \{H\}$ is H; this forces the curvature of $\mathbb{A}$ to have integral over $\Gamma_z \times S^1$ equal to $-2\pi\mathbf{i}$). Of course, this is a manifestation of the impossibility of extending the universal Jacobian to the whole Deligne–Mumford moduli space as a fibration of smooth orbifolds. See [Fr] for a more general discussion.

4. Orbibundles over orbisurfaces

We will call an **orbisurface** a smooth complex curve C together with a list of points $\mathbf{x} = x_1, \ldots, x_n \in C$ and corresponding positive integers $\mathbf{q} = q_1, \ldots, q_n$. Take an orbisurface $(C, \mathbf{x}, \mathbf{q})$. We now define (in a rather *ad hoc* way) what is an orbibundle over $(C, \mathbf{x}, \mathbf{q})$. (Of course, orbibundles are well known objects and can be defined in much greater generality than we do; we just recall the definition in this particular case to fix notations.) Pick a neighborhood U_j of each point x_j and a holomorphic covering map $\rho_j : \overline{U}_j \to U_j$ of degree q_j, such that $\overline{U}_j$ is biholomorphic to the unit disk and ρ_j has

maximal ramification at the unique point $\overline{x}_j := \rho_j^{-1}(x_j)$ and is unramified everywhere else. Let $\Gamma_j := \mathbb{Z}/q_j\mathbb{Z}$. The group Γ_j acts on $\overline{U}_j$ leaving the map ρ_j invariant. We will assume that the sets $U_1, \dots, U_n$ are disjoint. A **principal** S^1 **orbibundle** over $(C, \mathbf{x}, \mathbf{q})$ is a tuple

$$\mathbf{P} = (P, P_1, \dots, P_n, \psi_1, \dots, \psi_n),$$

where $P \to C \setminus \mathbf{x}$ is a principal S^1 bundle and, for each j, $P_j \to \overline{U}_j$ is a Γ_j-equivariant principal S^1 bundle and $\psi_j : \rho_j^* P|_{U_j \setminus \{x_j\}} \xrightarrow{\sim} P_j|_{\overline{U}_j \setminus \{\overline{x}_j\}}$ is an isomorphism. We leave to the reader the definition of isomorphism of principal orbibundles.

To define the **degree** of the orbibundle $\mathbf{P}$, pick a collection of local trivialisations $\tau = (\tau_1, \dots, \tau_n) \in T(P, \mathbf{x})$. Each of them gives rise (pulling back through ρ_j and using ψ_j) to a trivialisation $\overline{\tau}_j$ of P_j restricted to $\partial \overline{U}_j$. Such trivalisation allows to define a quotient of P_j over $\overline{U}_j / \partial \overline{U}_j \simeq S^2$, and we define $\deg(P_j, \overline{\tau}_j)$ to be the degree of this quotient. Then

$$\deg \mathbf{P} := \deg P^\tau - \sum \frac{1}{q_j} \deg(P_j, \overline{\tau}_j).$$

One checks, using (3.13), that this expression does not depend on the choice of τ.

4.1. **Relation to meromorphic connections.** Suppose that $(C, \mathbf{x})$ is a marked smooth curve and that $P \to C \setminus \mathbf{x}$ is a principal S^1 bundle. Let A be a meromorphic connection on P, all of whose residues are of the form $\mathbf{i}l$ for $l \in \mathbb{Q}$. Then we obtain in a natural way a structure of orbisurface on C and an orbibundle $\mathbf{P}^A$ on it which restricts to P on $C \setminus \mathbf{x}$, as follows. Pick a collection of trivialisations $\tau \in T(P, \mathbf{x})$ and let the residue $\mathrm{Res}(A, x_j, \tau_j)$ be $\mathbf{i}p_j/q_j$, where p_j and $q_j \geq 1$ are relatively prime integers. The resulting orbisurface is then $\mathbf{C}^A := (C, \mathbf{x}, \mathbf{q})$, where $\mathbf{q} = (q_1, \dots, q_n)$, and we define $P_j \to \overline{U}_j$ to be the trivial bundle with the trivial action of Γ_j. So it remains to define the isomorphisms ψ_j (which should be a trivialisation of $\rho_j^* P$, since P_j is the trivial bundle). The residue of the connection $\rho_j^* A$ on the bundle $\rho_j^* P$ at the point $\overline{x}_j$ with respect to $\rho_j^* \tau_j$ is $p_j \in \mathbb{Z}$. Then we take ψ_j to be the trivialisation of $\rho_j^* P$ with respect to which the covariant derivative $d_{\rho_j^* A}$ is equal to $d + \alpha$, where α extends continuously to $\Omega^1(\overline{U}_j, \mathbf{i}\mathbb{R})$ (such trivialisation exists by Lemma 3.1).

It is now a consequence of the Chern–Weil formula in Lemma 3.5 that

$$\deg \mathbf{P}^A = \frac{\mathbf{i}}{2\pi} \int_{C \setminus \mathbf{x}} F_A.$$

4.2. **Associated bundles.** Of course, one can consider in general orbibundles which are not necessarily principal S^1 orbibundles, but which are, more generally, **locally trivial orbibundles**. Their definition is completely analogous to that of S^1 principal orbibundle, substituting in each case the words "principal S^1 bundle" by "locally trivial bundle". The notion of **section** of an orbibundle is also almost evident; so, for example, if $\mathbf{Y} := (Y, \{Y_j\}, \{\psi_j^Y\})$ denotes a locally trivial orbibundle over an orbisurface $(C, \mathbf{x}, \mathbf{q})$, then a section of $\mathbf{Y}$ is a collection of sections $\Phi = (\phi, \phi_1, \dots, \phi_n)$, where $\phi \in \Gamma(C \setminus \mathbf{x}, Y)$ and $\{\phi_j \in \Gamma(\overline{U}_j, Y_j)\}$, satisfying the compatibility condition $\psi_j \circ \rho_j^* \phi = \phi_j$.

This is the main example which we will encounter. Given a principal S^1 orbibundle $\mathbf{P} = (P, \{P_j\}, \{\psi_j\})$ over $(C, \mathbf{x}, \mathbf{q})$, we can define a locally trivial orbibundle $\mathbf{Y} := \mathbf{P} \times_{S^1} X$ on $(C, \mathbf{x}, \mathbf{q})$ by setting $\mathbf{Y} := (Y, \{Y_j\}, \{\psi_j^Y\})$, where $Y = P \times_{S^1} X$, $Y_j = P_j \times_{S^1} X$ and ψ_j^Y denotes the isomorphism induced by ψ_j.

Recall that the de Rham complex of differential forms can be defined for orbifolds, and that its cohomology is isomorphic to the singular cohomology of the orbifold with real coefficients. This means in particular that we can use Chern–Weil theory to obtain representatives of Chern classes of orbibundles (allowing, in particular, to compute their degree). Another consequence is that one can define also the cohomology class $[\omega(A)]$ exactly as in Section 2.

5. Stable twisted holomorphic maps

5.1. **Critical residues.** Let $F \subset X$ be the fixed point set of the action of S^1. For each connected component $F' \subset F$, there is an action of S^1 on the corresponding normal bundle $N \to F'$. Then N splits as a direct sum $N = \bigoplus_{\chi \in \mathbb{Z}} N_\chi$, where $N_\chi \subset N$ is the subbundle on which S^1 acts with weight χ. Define the set of weights of F' to be $\mathrm{weight}(F') := \{\chi \in \mathbb{Z} \mid N_\chi \neq 0\}$. Define also the **set of weights of** X to be

$$\mathrm{weight}(X) := \bigcup_{F' \subset F} \mathrm{weight}(F') \subset \mathbb{Z},$$

where the union runs over the set of connected components of F. Finally, define the set of **critical residues** to be

$$\Lambda_{\mathrm{cr}} := \{\lambda \in \mathbf{i}\mathbb{R} \mid \text{there is some } w \in \mathrm{weight}(X) \text{ such that } w\lambda \in \mathbf{i}\mathbb{Z}\}.$$

For any $\lambda \in \mathbf{i}\mathbb{R}$ we will denote

$$X^\lambda := \{x \in X \mid e^{2\pi\lambda} \cdot x = x\}.$$

Of course, for any λ we have $F \subset X^\lambda$, and the condition $\lambda \in \Lambda_{\mathrm{cr}}$ is equivalent to the inclusion $F \subset X^\lambda$ being proper.

If A is a meromorphic connection on a principal S^1 bundle over a nodal curve C and $z \in C$ is a node, we will say that the **holonomy of** A **around** z **is critical** if, denoting by y any preimage of z in the normalisation of C, we have $H := \mathrm{Hol}(A, y) \in e^{2\pi\Lambda_{\mathrm{cr}}}$ (this is clearly independent of the chosen preimage of the normalisation, and hence well defined). In other words, the holonomy H is critical if the set of points in X fixed by H is bigger than F.

5.2. **Local behaviour of holomorphic sections near a marked point.** Let $(C, \mathbf{x})$ be a smooth marked curve and let P be a principal S^1 bundle over $C \setminus \mathbf{x}$, endowed with a meromorphic connection A. Let $Y := P \times_{S^1} X$, let $x \in \mathbf{x}$ be a marked point, and denote by $D_\epsilon \subset C$ the disk of radius ϵ centered at x. Following the same ideas as in Section 3.4 we can define an equivalence relation on the restriction of Y on D_ϵ using parallel transport with respect to A along radial directions. Taking the quotient by this

equivalence relation gives rise to a bundle Y_x over S_x, which can be canonically identified with $P_x \times_{S^1} X$.

Given a smooth section ϕ of Y, we say that ϕ **extends at x to give a section ϕ_x** of Y_x if, for every $\theta \in S_y$ we have $\lim_{\delta\to 0}[\phi(\exp_x \delta\theta)] = \phi_x(\theta)$, where the brackets denote the equivalence class in Y_x. In other words, this means that the pullback of the section ϕ to the real blowup of C at x extends to the exceptional divisor.

Theorem 5.1. *Suppose that a section ϕ of Y satisfies $\overline{\partial}_A\phi = 0$ and $\|d_A\phi\|_{L^2(C\setminus\mathbf{x})} < \infty$. Then for every marked point $x \in \mathbf{x}$ the section ϕ extends at x to give a section ϕ_x of Y_x and we have $d_{A_x}\phi_x = 0$. Furthermore, the section ϕ_x takes values in X^λ. In particular, if the holonomy of A around x is not critical, then ϕ_x takes values in the fixed point set F, hence ϕ_x is constant and the following limit exists*

$$\phi(x) := \lim_{z\to x} \phi(z) \in F. \tag{5.16}$$

The result of Theorem 5.1 is of local nature, so it follows from the corresponding version for the punctured disk, which is considered in Corollary 10.2 (see Section 10 for the statement and the proof of Corollary 10.2).

5.3. **Chains of gradient segments.** Recall that we denote $H := -\mathbf{i}\mu$. Let $\xi_t : X \to X$ be the downward gradient flow at time t of H, so $\xi_0 = \mathrm{Id}_X$ and

$$\frac{\partial \xi_s}{\partial t} = -\xi_s^*\nabla H = -\xi_s^* I\mathcal{X},$$

where $\mathcal{X}$ is the vector field generated by the infinitesimal action of $\mathbf{i} \in \mathrm{Lie}\, S^1$ on X. A **pointed gradient segment** in X is a pair (x, T), where x is a point in X, not contained in the fixed point set, and $T \subset \mathbb{R}$ is a closed interval of positive measure. Define $\xi_T(x) := \{\xi_t(x) \mid t \in T\} \subset X$. Two pointed gradient segments (x, T) and (x', T') are said to be **equivalent** if $\xi_T(x) = \xi_{T'}(x')$. A **gradient segment** in X is an equivalence class of pointed gradient segments. A **chain of gradient segments** is a finite sequence $\mathcal{T}$ of gradient segments represented by a list of pointed gradient segments $((x_1, T_1), \dots, (x_k, T_k))$ satisfying the following properties:

1. if $j > 1$ then $\inf T_j = -\infty$, and if $j < k$ then $\sup T_j = \infty$;
2. if $1 \le j < k$ then $\lim_{t\to\infty} \xi_t(x_j) = \lim_{l\to-\infty} \xi_l(x_{j+1})$.

The **beginnig** of $\mathcal{T}$ is the point $\lim_{l\to\inf T_1} \xi_l(x_1)$, and the **end** of $\mathcal{T}$ is $\lim_{t\to\sup T_k} \xi_t(x_k)$. A **degenerate chain of gradient segments** is simply a point $x \in X$ (so this corresponds to the case (x, T) where $T = \{0\}$).

Denote by $\mathcal{T}(X)$ the **set of chains of gradient segments on** X, including the degenerate ones. The group S^1 acts on $\mathcal{T}(X)$ as follows: if $\theta \in S^1$ and $\mathcal{T} \in \mathcal{T}(X)$ is represented by $((x_1, T_1), \dots, (x_k, T_k))$ then $\theta \cdot \mathcal{T}$ is represented by $((\theta \cdot x_1, T_1), \dots, (\theta \cdot x_k, T_k))$. The set $\mathcal{T}(X)$ carries a natural topology induced by the Hausdorff distance between subsets of X. With this topology, $\mathcal{T}(X)$ is clearly compact.

If S denotes the circle and $P \to S$ is a principal S^1 bundle provided with a connection A with trivial holonomy, then we say that a section $\mathcal{T}_S$ of the associated bundle $P \times_{S^1} \mathcal{T}(X)$

is **covariantly constant** (with respect to A) if, given a trivialisation with respect to which $d_A = d$, the map $S \to \mathcal{T}(X)$ given by the section $\mathcal{T}_S$ is constant.

5.4. **Metrics of fixed volume on stable curves.** Recall that the Deligne–Mumford moduli space $\overline{\mathcal{M}}_{g,n}$ of isomorphism classes of stable curves $[C, \mathbf{x}]$ admits a natural structure of orbifold. The map $f : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ which forgets the last point and stabilises gives $\overline{\mathcal{M}}_{g,n+1}$ the structure of universal curve over $\overline{\mathcal{M}}_{g,n}$. Let $\mathcal{M}et_{g,n}$ be the space of all smooth (in the orbifold sense) metrics in $\overline{\mathcal{M}}_{g,n+1}$ whose restriction to the fibre of f over $[C, \mathbf{x}]$ gives a metric on C of total volume 1 and in the conformal class defined by the complex structure. For any $\nu \in \mathcal{M}et_{g,n}$ and $[C, \mathbf{x}] \in \overline{\mathcal{M}}_{g,n}$ we will denote by $\nu_{[C,\mathbf{x}]}$ the induced metric in C. If $\pi : C' \to C$ is the normalisation map, the pullback $\pi^*\nu_{[C,\mathbf{x}]}$ is a smooth metric in C' (this is true because ν is smooth in $\overline{\mathcal{M}}_{g,n+1}$). Also, $\nu_{[C,\mathbf{x}]}$ is invariant under the action of the automorphisms of $(C, \mathbf{x})$. We give $\mathcal{M}et_{g,n}$ the obvious topology, which makes it a contractible space.

5.5. **Definition of c-stable twisted holomorhic maps.** Let g and n be nonnegative integers satisfying $2g+n \geq 3$. Take two natural numbers n_{cr} and n_{ge} such that $n_{\mathrm{cr}}+n_{\mathrm{ge}} = n$.

Let $(C, \mathbf{x})$ be a nodal curve of genus g and with n marked points. Repeatedly contracting the unstable components of C, we obtain a stable curve C^{st} and a map $s : C \to C^{\mathrm{st}}$, called the **stabilization map**. Let $C^b \subset C$ be the union of the irreducible components which are contracted to a point by the stabilisation map, and let $C^p \subset C$ be the union of the components not contained in C^b. Then we have $C = C^p \cup C^b$. The components of C^b are called the **bubble components** of C, and those of C^p the **principal components** of C. Finally, an **exceptional point** of C is a point which is either a marked point or a node.

Pick a metric $\nu \in \mathcal{M}et_{g,n}$ and an element c of $\mathbf{i}\mathbb{R}$. A **c-stable twisted holomorphic map** (c-STHM for short) of genus g and n marked points is a tuple

$$\mathcal{C} = ((C, \mathbf{x}_{\mathrm{cr}}, \mathbf{x}_{\mathrm{ge}}), (P, A, G), \phi, \{\mathcal{T}_y\}, \{\mathcal{T}_x\}),$$

where

(1) C is a connected compact nodal complex curve, $\mathbf{x}_{\mathrm{cr}}$ and $\mathbf{x}_{\mathrm{ge}}$ are disjoint lists of smooth points of C: $\mathbf{x}_{\mathrm{cr}}$ is the list of **critical marked points** and $\mathbf{x}_{\mathrm{ge}}$ is the list of **generic marked points**; $\mathbf{x}_{\mathrm{cr}}$ contains n_{cr} points and $\mathbf{x}_{\mathrm{ge}}$ contains n_{ge} points, and we denote by $\mathbf{x}$ the union $\mathbf{x}_{\mathrm{cr}} \cup \mathbf{x}_{\mathrm{ge}}$.

(2) P is a principal S^1 bundle on the set of non exceptional points of C,
$$P \to C \setminus (\mathbf{x} \cup \mathbf{z}),$$
where $\mathbf{z} \subset C$ is the set of nodes.

(3) A is a meromorphic connection on P, G is a choice of gluing data for P and ϕ is a section of the bundle $P \times_{S^1} X$.

(4) For each preimage y of a node in C, $\mathcal{T}_y$ is a covariantly constant section of the bundle $P_y \times_{S^1} \mathcal{T}(X^\lambda)$, where $\lambda \in \mathbf{i}\mathbb{R}$ is such that $\mathrm{Hol}(A, y) = e^{2\pi\lambda}$ (this gives a

chain of gradient lines for each tangent direction at y in the normalisation of the curve, varying in a S^1-equivariant way).

(5) For each $x \in \mathbf{x}_{\text{ge}}$, $\mathcal{T}_x$ is a covariantly constant section of the bundle $P_x \times_{S^1} \mathcal{T}(X^\lambda)$, where $\lambda \in \mathbf{i}\mathbb{R}$ is such that $\text{Hol}(A, x) = e^{2\pi\lambda}$.

The tuple $\mathcal{C}$ must satisfy the following conditions:

(1) **The section is holomorphic.** The section ϕ satisfies the equation
$$\overline{\partial}_A \phi = 0. \tag{5.17}$$
(2) **Vortex equation.** For any principal component $C_j \subset C^p$, let ν_j (resp. A_j, ϕ_j) be the restriction of $s^*\nu_{[C,\mathbf{x}]}$ (resp. A, ϕ) to C_j, where s is the stabilisation map, and let $d\,\text{vol}(\nu_j)$ be the induced volume form; then
$$\iota_{d\,\text{vol}(\nu_j)} F_{A_j} + \mu(\phi_j) = c; \tag{5.18}$$
we call this equation the vortex equation because in the case of $X = \mathbb{C}$ with the action of S^1 of weight 1 this equation coincides with the standard abelian vortex equation.

(3) **Flatness on bubbles.** The restriction of A to each bubble component is flat.

(4) **Finite energy.** The energy of ϕ as a section is bounded:
$$\|d_A \phi\|_{L^2} < \infty; \tag{5.19}$$
(5) **Matching condition at the nodes.** Let $z \in C$ be a node, and let y, y' be its preimages in the normalization map. By Theorem 5.1, (5.17) and (5.19) imply that ϕ extends to give sections of $\phi_y \in \Gamma(P_y \times_{S^1} X)$ and $\phi_{y'} \in \Gamma(P_{y'} \times_{S^1} X)$. Let $\rho : P_y \to P_{y'}$ be the isomorphism given by the gluing data G. Then
$$\mathcal{T}_y = \rho^* \mathcal{T}_{y'}$$
and for every $\theta \in S_y$, $\phi_y(\theta)$ has to be either the beginning or the end of the chain $\mathcal{T}_y(\theta)$, and $\rho^*\phi_{y'}(\theta)$ has to be the opposite extreme. Furthermore, if the holonomy of A around y is not critical, then the chain of gradient segments $\mathcal{T}_y$ has to be degenerate, which implies that $\rho^*\phi_{y'} = \phi_y$.

(6) **Matching condition and the generic marked points.** Given $x \in \mathbf{x}_{\text{ge}}$, let $\phi_x \in \Gamma(P_x \times_{S^1} X)$ be the extension of ϕ. For any $\theta \in S_x$, $\phi_x(\theta)$ is either the beginning or the end of $\mathcal{T}_x(\theta)$, and the opposite extreme of $\mathcal{T}_x(\theta)$ is a fixed point. Furthermore, if the holonomy of A around x is not critical then $\mathcal{T}_x$ has to be degenerate.

(7) **Stability condition for the bubbles.** If $C' \subset C$ is a bubble component with less than 3 exceptional points, then the restriction of $d_A\phi$ to C' is not identically zero.

Two c-STHM's $\mathcal{C}$ and $\mathcal{C}'$ are said to be **isomorphic** if there is a commuting diagram
$$\begin{tikzcd} P \arrow[r, "g"] \arrow[d] & P' \arrow[d] \\ C \arrow[r, "f"] & C', \end{tikzcd}$$

where g is an isomorphism of principal S^1 bundles and f is a biholomorphism satisfying $f(\mathbf{x}_{\mathrm{cr}}) = \mathbf{x}'_{\mathrm{cr}}$, $f(\mathbf{x}_{\mathrm{ge}}) = \mathbf{x}'_{\mathrm{ge}}$ and preserving the ordering of the marked points, such that

$$g^*A' = A, \qquad g^*G' = G, \qquad g^*\phi' = \phi, \qquad g^*\mathcal{T}'_y = \mathcal{T}_y \qquad \text{and} \qquad g^*\mathcal{T}'_x = \mathcal{T}_x.$$

5.6. **Remarks on the definition of Hamiltonian Gromov–Witten invariants.** We now make a few comments with the hope of clarifying some of the ingredients appearing in the definition of c-STHM. We recall that the main application we have in mind of the compactness result proved in this paper is the definition of invariants of the manifold X and the Hamiltonian action of S^1, using the moduli space of c-STHM's.

First of all, in this paper we have imposed no restriction on $c \in \mathbf{i}\mathbb{R}$. However, when considering the moduli space of c-STHM's we will need to take c away from a discrete set of critical values. The invariants obtained from two different choices of c may vary if we cross critical values when passing from one choice to the other (this is explained in [M], and is similar to the well known phenomenon of wall crossing in gauge theories). On the other hand, in order to define c-STHM's we have made a choice of an element $\nu \in \mathcal{M}et_{g,n}$; the invariants which we will construct do not depend on this choice, thanks to a standard cobordism argument and the fact the set $\mathcal{M}et_{g,n}$ is connected. This is analogous to the fact that Gromov–Witten invariants are independent of the chosen compatible almost complex structure. In our situation, the obtained invariants will also be independent of the S^1-invariant and compatible almost complex structure. Finally, the reason why we distinguish two different kinds of marked points (critical and generic) is the following: when constructing the moduli space, we will allow the residue at generic marked points to vary, whereas the residue at a critical marked point will always have to be critical. The way we define the evaluation map at a marked point will depend on whether the point is critical or generic.

6. Topology on the set of c-STHM

6.1. **Convergence of lines to chains of gradient lines.** Let $F \subset X$ be the fixed point set. For any small $\delta > 0$, denote by F^δ the δ-neighbourhood of F, that is, the set of points of X at distance $\leq \delta$ from F. Let also $X^\delta := X \setminus F^\delta$. Finally, recall that we denote by H the function $-\mathbf{i}\mu$.

Let $\{f_u : S_u \to X\}$ be a sequence of smooth maps, where each $S_u \subset \mathbb{R}$ is a closed interval. We call the pairs (f_u, S_u) **lines in** X. Let $\mathcal{T}$ be a chain of gradient flow lines (see Section 5.3) represented by a list $((x_1, T_1), \dots, (x_k, T_k))$ of pointed gradient segments. Let $F_1, \dots, F_l$ be the connected components of the fixed point set which intersect the closure of $\bigcup \xi_{T_k}(x_k)$, labelled in such a way that $H(F_1) > H(F_2) > \cdots > H(F_l)$. We will say that **the sequence of lines** $\{(f_u, S_u)\}$ **converges to** $\mathcal{T}$ if for each small enough $\delta > 0$ and any big enough u we can write S_u as a union of two sets

$$S_u = T_u^\delta \cup E_u^\delta, \tag{6.20}$$

in such a way that:

(1) each E_u^δ is a union of l closed intervals: $E_u^\delta = E_{u,1}^\delta \cup \dots \cup E_{u,l}^\delta$, and each T_u^δ is a union of k closed intervals: $T_u^\delta = T_{u,1}^\delta \cup \dots \cup T_{u,k}^\delta$, labelled in such a way that $\sup E_{u,j-1}^\delta < \inf E_{u,j}^\delta$ and $\sup T_{u,j-1}^\delta \leq \inf T_{u,j}^\delta$ for every j for which the expression makes sense;

(2) denoting by $t_{u,j}^\delta$ any point of $T_{u,j}^\delta$, the maps $f_u : T_{u,j}^\delta \to X$ approximate gradient segments

$$\lim_{u\to\infty} \sup_{t\in T_{u,j}^\delta} d(\xi_{t-t_{u,j}^\delta}(f_u(t_{u,j}^\delta)), f_u(t)) = 0; \tag{6.21}$$

(3) the images $f_u(T_{u,j}^\delta)$ approximate in X^δ the j-th gradient segment of $\mathcal{T}$:

$$\lim_{\delta\to 0} \lim_{u\to\infty} D(f_u(T_{u,j}^\delta, \xi_{T_j}(x_j) \cap X^\delta) = 0,$$

where here D denotes de Hausdorff distance between sets;

(4) the images of the sets $E_{u,j}^\delta$ become smaller and smaller as $\delta \to 0$ and they accumulate near the fixed point set:

$$\lim_{\delta\to 0}\left(\limsup_{u\to\infty} \operatorname{diam}(f_u(E_{u,j}^\delta))\right) = \lim_{\delta\to 0}\left(\limsup_{u\to\infty} d(f_u(E_{u,j}^\delta), F_j)\right) = 0.$$

This implies in particular that the sets $f_u(S_u)$ converge in the Hausdorff metric to $\bigcup \xi_{T_j}(x_j)$.

6.2. **Convergence with gauge λ of cylinders to chains of gradient lines.** Now suppose that $\{\phi_u : C_u \to X\}$ is a sequence of smooth maps, where each C_u is a cylinder $C_u = S_u \times S^1$ and $S_u \subset \mathbb{R}$ is a closed interval. We call the pairs (ϕ_u, C_u) **cylinders in** X. Let $\lambda \in \Lambda_{\mathrm{cr}}$ be a critical residue. We say that **the sequence of cylinders** $\{\phi_u, C_u\}$ **converges with gauge λ to a chain of gradient lines** $\mathcal{T}$ in X^λ if there is a sequence of lines in X, $\{(\psi_u, S_u)\}$, which converges to $\mathcal{T}$ and such that

$$\lim_{u\to\infty} \sup_{(t,\theta)\in C_u} d(e^{\lambda\theta}\phi_u(t,\theta), \psi_u(t)) = 0.$$

Here $e^{\lambda\theta}$ denotes *any* number of the form $e^{\lambda\overline{\theta}}$, where $\overline{\theta} \in \mathbb{R}$ is a lift of $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. Since the chain $\mathcal{T}$ is contained in X^λ, the resulting notion of convergence is independent of the chosen lifts.

6.3. **Description of the topology of $\overline{\mathcal{M}}_{g,n}$.** We give here a description of the topology of the Deligne–Mumford moduli space which is suitable for our purposes. For that it suffices to specify what it means that a sequence of stable (pointed) curves converges to a given curve.

Let $(C, \mathbf{x})$ be a stable curve, denote by $\mathbf{z}$ the set of nodes, let $\pi : C' \to C$ be the normalization and let $\mathbf{y}$ be the preimages of $\mathbf{z}$ by π. Pick a conformal metric on C' and let $\epsilon > 0$ be a small number. For each $y \in \mathbf{y}$, take a neighborhood $U_y \subset C'$, small enough so that it admits a biholomorphism $\zeta_y : D(\epsilon) \to U_y$ with the disk $D(\epsilon) \subset T_yC'$ centered at 0 and of radius ϵ. Chose also a small neighborhood $B_x \subset C'$ of the preimage of each

marked point $x \in \mathbf{x}$. Denote by $I \in \operatorname{End} TC'$ the complex structure of C', which we now view as a compact real surface.

Let now $I' \in \operatorname{End} TC'$ be another complex structure which coincides with I on each U_y and each B_x (call such complex structure **admissible**). Take also, for each node z with preimages y, y' in C', an element $\delta_z \in T_yC' \otimes T_{y'}C'$ satisfying $|\delta| < \epsilon^2$. We call the collection of numbers $\{\delta_z\}$ **smoothing parameters**. Define a new curve $C(I', \{\delta_z\})$ as follows: replace the complex structure I by I', then remove for each for each pair y, y' of preimages of a node z, the sets $\zeta_y(D(|\delta_z|/\epsilon))$ and $\zeta_{y'}(D(|\delta_z|/\epsilon))$ from C', and finally identify for each pair of elements $u \in D(\epsilon) \setminus D(\delta_z/\epsilon)$ and $v \in D(\epsilon) \setminus D(\delta_z/\epsilon)$ satisfying $u \otimes v = \delta_z$, the images $\zeta_y(u)$ and $\zeta_{y'}(v)$ (such identifications preserve the complex structure because I' is admissible).

For later use, define for every z the subset $N_y(\delta_z) \subset C(I', \{\delta_z\})$ to be the image by ζ_y of the annulus $D(\epsilon) \setminus D(|\delta_z|/\epsilon)$ (this is equal to the image by $\zeta_{y'}$ of the corresponding subset of $T_{y'}C'$). The set N_z is conformally equivalent to the cylinder: $[\ln|\delta_z| - \ln\epsilon, \ln\epsilon] \times S^1$. We will say that y **is in the side of** $\{\ln\epsilon\} \times S^1$, and that y' is in the side of $\{\ln|\delta_j| - \ln\epsilon\} \times S^1$ (if we consider $N_{y'}(\delta_z)$ instead, then the roles are inverted).

Note that for any compact set $K \subset C \setminus (\mathbf{x} \cup \mathbf{z})$ and small enough smoothing data $\{\delta_z\}$ there is a canonical inclusion $K \to C(I', \{\delta_z\})$, which we will denote by ι_K. (For ι_K to exist it suffices to take each δ_z so that $\zeta_y(D(\delta_z/\epsilon))$ and $\zeta_{y'}(D(\delta_z/\epsilon))$ are disjoint from K.)

A sequence $\{(C_u, \mathbf{x}_u)\}$ of stable curves **converges** to $(C, \mathbf{x})$ if for big enough u there is an admissible complex structure I_u, smoothing parameters $\{\delta_{u,z}\}$, and an isomorphism of marked nodal curves

$$\xi_u : (C(I_u, \{\delta_{u,z}\}), \mathbf{x}) \to (C_u, \mathbf{x}_u), \tag{6.22}$$

such that I_u converges to I in $C^\infty(\operatorname{End} TC')$ and for each node z we have $\delta_{u,z} \to 0$.

The topology on $\overline{\mathcal{M}}_{g,n}$ defined by this notion of convergence coincides with the usual one (see for example Section 9 in [FO], where the topology of $\overline{\mathcal{M}}_{g,h}$ is described in similar terms).

6.4. **Connections in balanced temporal gauge.** Let $C = [p, q] \times S^1$ be a cylinder, and denote by (t, θ) the usual coordinates. Let $d_A = d + \alpha$ be a covariant derivative on the trivial principal S^1 bundle over C. We will say that α is in **balanced temporal gauge** if it is in temporal gauge, so that $\alpha = a d\theta$ for some function $a : C \to \mathbf{i}\mathbb{R}$, and furthermore the restriction of a to the middle circle $\{(p+q)/2\} \times S^1$ is constantly equal to some $\lambda \in \mathbf{i}\mathbb{R}$, which is called the **residue** of A (with respect to the trivialization). Any connection on the trivial bundle over C is gauge equivalent to a connection in balanced temporal gauge. Furthermore, since $d\alpha = \frac{\partial a}{\partial t} dt \wedge d\theta$, we have the estimate:

$$|a(t, \theta) - \lambda| \le \left| \int_{(p+q)/2}^{t} |d\alpha(\tau, \theta)| d\tau \right|. \tag{6.23}$$

6.5. **Convergence of c-STHM.** Our aim here is to define a topology on the set of isomorphism classes of c-STHM's specifying as before what it means for a sequence of

c-STHM's to converge to a given c-STHM's. Before defining the convergence of sequences, we make the observation that the notion of convergence for stable curves given in Section 6.3 makes perfect sense when considering nodal marked curves in general: if $(C, \mathbf{x})$ is a nodal marked curve with k nodes, we can define as before the deformations $C(I_u, \delta_1, \dots, \delta_k)$. (Of course, the topology induced by this notion on the set of isomorphism classes of nodal marked curves is not Hausdorff.)

For simplicity, we will only define convergence of sequences of c-STHM's with smooth underlying marked curve and with degenerate chains of gradient segments at generic marked points. To pass from this to the general case is routine.

So let $\{\mathcal{C}_u\}$ be a sequence of c-SHTC's. Suppose that $(C_u, \mathbf{x}_u)$ is the smooth marked curve underlying $\mathcal{C}_u$, and that $\mathbf{x}_u = \mathbf{x}_{\mathrm{cr},u} \cup \mathbf{x}_{\mathrm{ge},u}$, (P_u, A_u) is the bundle and connection on C_u (since C_u is smooth there is no gluing data) and ϕ_u is the section of $P_u \times_{S^1} X$ (again, since C_u is smooth there are no chains of gradient segments $\mathcal{T}_y$); finally, for each $x \in \mathbf{x}_{\mathrm{ge}}$ the chain $\mathcal{T}_x$ is degenerate. Let now

$$\mathcal{C} = ((C, \mathbf{x}_{\mathrm{cr}}, \mathbf{x}_{\mathrm{ge}}), (P, A, G), \phi, \{\mathcal{T}_y\}, \{\mathcal{T}_x\})$$

be another c-STHM, not necessarily with smooth underlying curve, and let $\mathbf{z} \subset C$ be the nodes of C.

We will say that the sequence of isomorphism classes $\{[\mathcal{C}_u]\}$ converges to $[\mathcal{C}]$ if for any exhaustion $K_1 \subset \dots \subset K_l \subset \dots$ of $C \setminus (\mathbf{x} \cup \mathbf{z})$ by compact subsets the following holds.

(1) **Convergence of the underlying curves.** The curves $(C_u, \mathbf{x}_{\mathrm{cr},u}, \mathbf{x}_{\mathrm{ge},u})$ converge to $(C, \mathbf{x}_{\mathrm{cr}}, \mathbf{x}_{\mathrm{ge}})$. This implies that there are isomorphisms

$$\xi_u : (C(I_u, \{\delta_{u,z}\}), \mathbf{x}_{\mathrm{cr}}, \mathbf{x}_{\mathrm{ge}}) \to (C_u, \mathbf{x}_{\mathrm{cr},u}, \mathbf{x}_{\mathrm{ge},u})$$

such that $I_u \to I$ and $\delta_{u,z} \to 0$. Pulling back everything by ξ_u we can assume that the underlying curve of $\mathcal{C}_u$ is $C(I_u, \{\delta_{u,z}\}, \mathbf{x}_{\mathrm{cr}}, \mathbf{x}_{\mathrm{ge}})$.

(2) **Convergence of gluing angles.** Let $((\gamma_1, \rho_1), \dots, (\gamma_k, \rho_k))$ be the glueing data at each of the nodes of C given by G. Since we assume that each C_u is smooth, any gluing parameter $\delta_{u,j}$ is nonzero and hence gives rise to a gluing angle $[\delta_{u,j}] \in \Gamma_{z_j}$. Then, for any j we must have $[\delta_{u,j}] \to \gamma_j$.

(3) **Convergence of the connections and sections away from the nodes.** For each K_l and any big enough u (so that ι_{K_l} is defined) there must exist an isomorphism of vector bundles

$$\rho_{u,l} : P|_{K_l} \to \iota_{K_l}^* P_u$$

such that $\rho_{u,l}^* \iota_K^* A_u$ converges to A and $\rho_{u,l}^* \iota_K^* \phi_u$ converges to ϕ on K_l as u goes to ∞ (here the convergence is assumed to be in C^∞). This implies that the holonomies of A_u around the marked points converge to those of A.

(4) **Convergence of gluing data.** The isomorphisms $\rho_{u,l}$ have to satisfy the following additional condition. Suppose that $r > 0$ is smaller than the ϵ used in the definition of convergence of stable curves. Take some node $z \in \mathbf{z}$ with preimages y, y'. Let $Y(r) := \zeta_y(S(r))$, where $S(r) \subset T_y C'$ is the circle of radius r centered at 0, and define $Y'(r)$ similarly. Suppose that l is big enough so that both $Y(r)$ and

$Y'(r)$ are contained in K_l, and suppose that u is big enough so that ι_{K_l} exists. Define a map $\tau_{u,l}(r)$ by the condition that the following diagram commutes:

$$\begin{tikzcd} (\iota_{K_l}^* P_u)|_{Y(r)} \arrow[r, "\tau_{A_u}"] & (\iota_{K_l}^* P_u)|_{Y'(r)} \\ P|_{Y(r)} \arrow[u, "\rho_{u,z}"] \arrow[r, "\tau_{u,l}(r)"] & P|_{Y'(r)}, \arrow[u, "\rho_{u,l}"'] \end{tikzcd}$$

where τ_{A_u} denotes the parallel transport along the images by ζ_y of the radial directions in the annulus $D(\epsilon) \setminus D(\delta_z/\epsilon)$. Finally, let $f' : P|_{Y'(r)} \to P_{y'}$ and $f : P|_{Y(r)} \to P_y$ be the natural projections. Then the composition $f' \circ \tau_{u,l}(r) \circ f$ gives an isomorphism between P_y and $P_{y'}$ lifting the isometry between S_y and $S_{y'}$ given by the gluing angle $[\delta_z]$, so it specifies an element $g_{u,l,j}(r) \in \Gamma(P, z)$. The condition is that for every j

$$\lim_{r\to 0} \left(\limsup_{u\to\infty} \limsup_{l\to\infty} d(g_{u,l,j}(r), (\gamma_j, \rho_j)) \right) = 0,$$

where d denotes a fixed distance function defined on $\Gamma(P, z_j)$.

(5) **Convergence near the nodes to chains of gradient segments.** Fix some node $z \in C$ with preimages y, y' in the normalization. Suppose that the holonomy of A around y is $e^{2\pi\lambda}$ for some $\lambda \in \mathbf{i}\mathbb{R}$. We will denote $\delta_u := \delta_{u,z}$. To specify the condition, we will use the cylinders

$$N_u := N_y(\delta_u) \simeq [\ln|\delta_u| - \ln\epsilon, \ln\epsilon] \times S^1$$

defined in Section 6.3. Define, for any big $\Delta > 0$ and u the cylinder

$$N_u(\Delta) = [\ln|\delta_u| - \ln\epsilon + \Delta, \ln\epsilon - \Delta] \times S^1 \subset N_u.$$

We distinguish two situations.

(a) If the chain of gradient segments $\mathcal{T}_y$ is degenerate (for example, if the residue λ is not critical), then we must have

$$\lim_{\Delta\to 0} \left(\limsup_{u\to\infty} \operatorname{diam}(\phi_u(N_u(\Delta)) \right) = 0.$$

(b) Otherwise, the following must happen; for each big enough u there is a cylinder $M_u \subset N_u$ satisfying two properties. First, for any $\Delta > 0$ and big enough u, we must have $M_u \subset N_u(\Delta)$. Then the complementary is the union of two cylinders $M_u^{\Delta,-}$ and $M_u^{\Delta,+}$. The second property is the following. Suppose that P_u is trivialized on $N(\delta_u)$ in such a way that $d_{A_u} = d + \alpha_u$ is in balanced temporal gauge and with residue λ_u, and that $\lambda_u \to \lambda$. Then the section ϕ_u restricts to give a map $\phi_u : N_u \to X$, and we must have:

(i) the images of the sets $M_u^{\delta,\pm}$ have smaller and smaller diameter:

$$\lim_{\Delta\to\infty} \left(\limsup_{u\to\infty} \operatorname{diam}_{S^1}(\phi_u(M_u^{\Delta,\pm})) \right) = 0;$$

(ii) $\lambda_u \neq \lambda$ for big enough u, so the number $l_u := -\mathbf{i}(\lambda_u - \lambda)$ is nonzero;

(iii) the sequence of cylinders $\{\phi_u/l_u : l_uM_u \to X\}$ converges with gauge λ to the chain of gradient segments $\mathcal{T}_{y_j}$ (here, if $M_u = [p,q] \times S^1$ then $l_uM_u = [l_up, l_uq]$ and $\phi_u/l_u(t,\theta) := \phi_u(t/l_u, \theta)$).

(6) **Convergence near generic marked points to chains of gradient segments.** This is similar to the previous condition. Let $x \in \mathbf{x}_{\mathrm{ge}}$. Let N be a punctured neighborhood of x biholomorphic to $[0,\infty) \times S^1$. Define also $N(\Delta) := [\Delta, \infty)$. We distinguish two cases.

(a) If the chain $\mathcal{T}_x$ is degenerate then

$$\lim_{\Delta\to 0}\left(\limsup_{u\to\infty} \operatorname{diam}(\phi_u(N(\Delta))\right) = 0.$$

(b) Suppose now that $\mathcal{T}_x$ is not degenerate. Let the holonomy of A around x be $e^{2\pi\lambda}$. Then we must have $\lambda \in \Lambda_{\mathrm{cr}}$. There must be, for each big enough u, a cylinder $M_u = [\Delta_u, \infty)$ satisfying:

(i)

$$\lim_{\Delta\to\infty}\left(\limsup_{u\to 0} \operatorname{diam}_{S^1}(\phi_u([\Delta, \Delta_u] \times S^1))\right) = 0;$$

(ii) take a trivialisation of P_u with respect to which $d_{A_u} = d + \alpha_u$ is in temporal gauge and the restriction of α_u to $\{\Delta_u\} \times S^1$ is equal to $\lambda_u d\theta$, where $\lambda_u \in \mathbf{i}\mathbb{R}$ is a constant, and $\lambda_u \to \lambda$. We must have $\lambda_u \neq \lambda$ for big enough u, so the number $l_u := -\mathbf{i}(\lambda_u - \lambda)$ is nonzero;

(iii) the sequence of cylinders $\{\phi_u/l_u : l_uM_u \to X\}$ converges with gauge λ to the chain of gradient segments $\mathcal{T}_x$.

7. The Yang–Mills–Higgs functional $\mathcal{YMH}_c$

Let $\mathcal{C} = ((C, \mathbf{x}_{\mathrm{cr}}, \mathbf{x}_{\mathrm{ge}}), (P, A, G), \phi, \{\mathcal{T}_y\}, \{\mathcal{T}_x\})$ be a c-STHM. Let C^p (resp. C^b) be the union of the principal components (resp. bubble components) of C. Let $s : C \to C^{\mathrm{st}}$ be the stabilization map. Take on C^p the conformal metric $s^*\nu_{[C^{\mathrm{st}},\mathbf{x}]}$ and extend it to a conformal metric on C by chosing an arbitrary conformal metric on the bubbles. Define

$$\mathcal{YMH}_c(\mathcal{C}) := \|F_A\|^2_{L^2(C^p)} + \|\mu(\phi) - c\|^2_{L^2(C^p)} + \|d_A\phi\|^2_{L^2(C)}.$$

We call $\mathcal{YMH}_c$ the **Yang–Mills–Higgs functional**. Its value is independent of the conformal metric chosen in the bubbles, since over the bubbles we only integrate the energy $\|d_A\phi\|^2_{L^2(C)}$, which is conformaly invariant. In this section we will compute $\mathcal{YMH}_c(\mathcal{C})$ in terms of topological data and the residues of the meromorphic connection A.

7.1. **Orbifold structure induced by the critical residues.** Let $(C, \mathbf{x})$ be a smooth marked curve, let $P \to C \setminus \mathbf{x}$ be a principal S^1 bundle and let A be a meromorphic conection on P. We can divide the set of marked points as $\mathbf{x} = \mathbf{x}_{\mathrm{cr},A} \cup \mathbf{x}_{\mathrm{ge},A}$, where $\mathbf{x}_{\mathrm{cr},A}$ (resp. $\mathbf{x}_{\mathrm{ge},A}$) denotes the set of points around which the holonomy of A is critical (resp. noncritical, which we call generic). Since critical residues are of the form $\mathbf{i}l$ with $l \in \mathbb{Q}$, we can use the construction in Section 4.1 to obtain, for any $\tau \in T(P, \mathbf{x}_{\mathrm{cr},A})$, an orbibundle $\mathbf{P}^{\tau,A} = (P^\tau, \{P_j\}, \{\psi_j\})$ on $\mathbf{C}^A := (C, \mathbf{x}_{\mathrm{cr},A}, \mathbf{q})$ (where $\mathbf{q}$ denotes the collection

of denominators in the set of critical residues of A). The following result gives a crucial property of holomorphic sections of $P \times_{S^1} X$ with bounded energy.

Theorem 7.1. *Suppose that ϕ is a section of $P \times_{S^1} X$ which satisfies $\overline{\partial}_A \phi = 0$ and $\|d_A\phi\|_{L^2(C \setminus \mathbf{x})} < \infty$. Then, for any $\tau \in T(P, \mathbf{x}_{\mathrm{ge},A})$ the section ϕ extends to give a section $\Phi^{\tau,A}$ of the orbibundle $\mathbf{P}^{\tau,A}$.*

Proof. The fact that ϕ extends to the points in $\mathbf{x}_{\mathrm{ge},A}$ follows from Theorem 5.1 (since in this case the limiting orbit has to be a fixed point, there is indeed an extension no matter what trivialisation τ we chose). The fact that ϕ extends to $\mathbf{x}_{\mathrm{cr},A}$ is a consequence of Gromov removal of singularities theorem. Indeed, pulling back (A, ϕ) to $\overline{U}_j \setminus \{\overline{x}_j\}$ and applying a suitable gauge transformation g the connection extends to the whole $\overline{U}_j$, hence defines a complex structure on $\overline{U}_j \times X$. The resulting map $g \cdot \rho_j^* \phi : \overline{U}_j \setminus \{\overline{x}_j\} \to \overline{U}_j \times X$ is complex and has bounded energy, so Gromov's theorem applies (see for more detais the proof the case of critical weight of Theorem 10.1 in Section 10.4). □

7.2. Computation of the Yang–Mills–Higgs functional.

Lemma 7.2. *Let $(C, \mathbf{x})$ be a smooth marked curve, let $P \to C \setminus \mathbf{x}$ be an principal S^1 bundle, let A be a meromorphic connection on P and let ϕ be a section of $P \times_{S^1} X$. Suppose that $\overline{\partial}_A \phi = 0$ and that $\|d_A\phi\|_{L^2} < \infty$. Then, for any $\tau \in T(P, \mathbf{x}_{\mathrm{ge},A})$, we have*

$$\int (\Phi^{\tau,A})^*[\omega(A)] = \frac{1}{2}\|d_A\phi\|_{L^2}^2 - \int F_A \mu(\phi) - \sum_{x \in \mathbf{x}_{\mathrm{ge},A}} 2\pi \operatorname{Res}(A, x, \tau_x)\mu(\phi(x)).$$

Remark 7.3. *Note that, since A might have nontrivial poles, the form $\omega(A)$ is singular on $\mathbf{P}^{\tau,A} \times_{S^1} X$. However, one can define the cohomology class $[\omega(A)]$ simply by taking any smooth connection A' and setting $[\omega(A)] := [\omega(A')]$. By Remark 2.2, this is independent of A'.*

Proof. The proof is similar to that of Chern–Weil formula for meromorphic connections in Lemma 3.5. For any small $\epsilon > 0$ we modify A near each point $x \in \mathbf{x}_{\mathrm{ge},A}$ as follows: if in a neighbourhood of x we have $d_A = d + \alpha + \lambda d\theta$ (where α extends continuously to x and $\lambda = \operatorname{Res}(A, x, \tau_x)$), then we replace $\lambda d\theta$ by η_ϵ, where η_ϵ extends smoothly to x, it coincides with $\lambda d\theta$ away from the disk $B(x, \epsilon)$ of radius ϵ centered at x, and both the supremum of the norm $|\eta_\epsilon|$ and the integral of $|d\eta_\epsilon|$ on $B(x, \epsilon)$ is bounded above by a constant C independent of ϵ (for example, one can define η_ϵ to be $r^2 \lambda d\theta / \epsilon^2$ in $B(x, \epsilon - \epsilon^2)$). Let A_ϵ be the resulting smooth connection on $\mathbf{P} := \mathbf{P}^{\tau,A}$ over the orbisurface $\mathbf{C} = \mathbf{C}^A$. Then we have for any $\epsilon > 0$

$$\int (\Phi^{\tau,A})^*[\omega(A)] = \int (\Phi^{\tau,A})^*[\omega(A_\epsilon)] = \frac{1}{2}\|d_{A_\epsilon}\phi\|_{L^2}^2 - \int F_{A_\epsilon}\mu(\phi).$$

This follows from formula (2.10) and the fact that (taking local trivialisations for (2.10) to make sense) for any point $x \in C$ and any choice of conformal metric on C we have

$$\phi(x)^*(\pi_X^*\omega - \pi^*\alpha_\epsilon \wedge \pi_X^* \iota_{\mathcal{X}}\omega) = \frac{1}{2}(|\partial_{A_\epsilon}\phi|^2 - |\overline{\partial}_{A_\epsilon}\phi|^2)\, dvol(x).$$

To finish the proof, we make two observations. First, that as $\epsilon \to 0$, $\|d_{A_\epsilon}\phi\|^2_{L^2}$ converges to $\|d_A\phi\|^2_{L^2}$. This is a consequence of the bound $|\eta_\epsilon| < C$ and of Theorem 10.1. Second, we similarly have, as $\epsilon \to 0$,

$$\int F_{A_\epsilon}\mu(\phi) \to \int F_A\mu(\phi) + \sum_{x\in\mathbf{x}_{\mathrm{ge},A}} 2\pi\,\mathrm{Res}(A,x,\tau_x).$$

Indeed, setting $C_\epsilon := C \setminus \bigcup_{x\in\mathbf{x}_{\mathrm{ge},A}} B(x,\epsilon)$ we have

$$\int F_{A_\epsilon}\mu(\phi) = \int_{C_\epsilon} F_A\mu(\phi) + \sum_{x\in\mathbf{x}_{\mathrm{ge},A}} \int_{B(x,\epsilon)} F_{A_\epsilon}\mu(\phi).$$

The first integral on the right hand side clearly converges to $\int_C F_A\mu(\phi)$. To estimate the second term we compute for any $x \in \mathbf{x}_{\mathrm{ge},A}$

$$\int_{B(x,\epsilon)} F_{A_\epsilon}\mu(\phi) = \int_{z\in B(x,\epsilon)} F_{A_\epsilon}(z)\mu(\phi(x))dz + \int_{z\in B(x,\epsilon)} F_{A_\epsilon}(z)(\mu(\phi(z)) - \mu(\phi(x)))dz.$$

The last integral converges to 0 as $\epsilon \to 0$ because $\int_{B(x,\epsilon)} |d\eta_\epsilon| < C$ and because $\mu(z) \to \mu(x)$ as $z \to x$ (this follows from Corollary 10.2). Finally, we compute

$$\begin{aligned}\int_{z\in B(x,\epsilon)} F_{A_\epsilon}(z)\mu(\phi(x))dz &= \int_{z\in B(x,\epsilon)} d\eta_\epsilon\mu(\phi(x))dz + \int_{z\in B(x,\epsilon)} d\alpha\mu(\phi(x))dz \\ &= \int_{\partial B(x,\epsilon)} \eta_\epsilon\mu(\phi(x))dz + \int_{z\in B(x,\epsilon)} d\alpha\mu(\phi(x))dz \\ &= 2\pi\,\mathrm{Res}(A,x,\tau_x)\mu(\phi(x)) + \int_{z\in B(x,\epsilon)} d\alpha\mu(\phi(x))dz,\end{aligned}$$

and the last integral converges to 0 because $d\alpha$ is integrable. □

Theorem 7.4. *Let $\mathcal{C} = ((C,\mathbf{x}_{\mathrm{cr}},\mathbf{x}_{\mathrm{ge}}),(P,A,G),\phi,\{\mathcal{T}_y\},\{\mathcal{T}_x\})$ be a c-STHM. We have $\mathbf{x}_{\mathrm{ge},A} \subset \mathbf{x}_{\mathrm{ge}}$ and, for any $\tau \in T(P,\mathbf{x}_{\mathrm{ge},A})$,*

$$\mathcal{YMH}_c(\mathcal{C}) = 2\int(\Phi^{\tau,A})^*[\omega(A)] - 4\pi\mathbf{i}c\deg\mathbf{P}^{\tau,A} + 4\pi\sum_{x\in\mathbf{x}_{\mathrm{ge}}}\mathrm{Res}(A,x,\tau_x)(\mu(\phi(x)) - c).$$

Proof. The inclusion $\mathbf{x}_{\mathrm{ge},A} \subset \mathbf{x}_{\mathrm{ge}}$ follows from the definition of c-STHM's. So we only have to prove the formula for $\mathcal{YMH}_c(\mathcal{C})$. For that we can compute separatedly the integral on each principal component $C_i \subset C^p$ (the computation for the bubbles is as in standard Gromov–Witten theory, except the case in which the connection is flat but has poles in two points of the bubble, which follows from Lemma 7.2). Using Lemma 7.2 and the Chern–Weil formula in Lemma 3.5 we have (here all integrals are over C_i, $\mathbf{y}_{\mathrm{ge}}$ denotes the marked points and nodes of C_i on which A has a pole with noncritical

residue, v denotes the volume form, and all norms are L^2 over C_i):

$$\begin{aligned}
0 &= \|\iota_v F_A + \mu(\phi) - c\|^2 \\
&= \|F_A\|^2 + \|\mu(\phi) - c\|^2 - 2\int F_A c + 2\int F_A \mu(\phi) \\
&= \|F_A\|^2 + \|\mu(\phi) - c\|^2 + 4\pi \mathbf{i} c \left(\deg \mathbf{P}^{\tau,A} - \mathbf{i} \sum_{x \in \mathbf{y}_{\mathrm{ge}}} \operatorname{Res}(A, x, \tau_x) \right) \\
&\quad + \|d_A \phi\|^2 - 4\pi \sum_{x \in \mathbf{y}_{\mathrm{ge}}} \operatorname{Res}(A, x, \tau_x) \mu(\phi(x)) - 2\int (\Phi^{\tau,A})^* [\omega(A)].
\end{aligned}$$

Rearranging the terms we obtain the desired formula. $\square$

8. Bounding the number of bubbles in terms of $\mathcal{YMH}_c$

Theorem 8.1. *For any $K > 0$ and any g, n satisfying $2g + n \geq 3$ there exists some N with the following property. Let $\mathcal{C} = ((C, \mathbf{x}), (P, A, G), \phi, \{\mathcal{T}_y\}, \{\mathcal{T}_x\})$ be a c-STHM of genus g and n marked points. Suppose that $\mathcal{YMH}_c(\mathcal{C}) \leq K$. Then the number of bubbles in C is less that N.*

The proof of Theorem 8.1 is given in Section 8.3. Note that the corresponding result in Gromov–Witten theory is an immediate consequence of the existence of a lower bound on the energy of nontrivial bubbles. In our situation such a lower bound does not exist, and this is why the theorem is not so obvious. The idea of the proof is the observation that bubbles with very little energy come in fact from gradient segments, and that a chain of consecutive bubbles with little energy gives rise to a chain of gradient segments, with at least as many components as the chain of bubbles. The number of components in a chain of gradient segments is at most equal to the number of connected components of the fixed point set of X. Hence, there is an upper bound for the length of a chain of consecutive bubbles with little energy.

8.1. Connecting bubbles and tree bubbles.

Take some c-STHM $\mathcal{C}$, let $(C, \mathbf{x})$ be the corresponding marked curve, let P be the principal bundle, and let A be the meromorphic connection. Let $\Gamma := \Gamma(C, \mathbf{x})$ be the graph whose set of vertices is the set of irreducible components of C plus the set of marked points, and whose set of edges is the following:

- for any pair of vertices v', v'' corresponding to components C', C'', there are as many edges connecting v' to v'' as there are nodes in C at which C' and C'' meet (in particular, Γ has a loop for each node whose two branches belong to the same component);
- if v_x is a vertex corresponding to a marked point $x \in \mathbf{x}$, then there is an edge connecting v_x to the vertex corresponding to the component of C in which x is contained.

We divide the set of vertices of Γ in **bubble vertices**, **principal vertices** and **marked point vertices**. A bubble vertex v of Γ is said to be **exterior** if there is a unique edge

having v as one of its extremes (such edge is not allowed to be a loop). A bubble vertex v of Γ is called a **tree vertex** if there exists a saturated subgraph $T \subset \Gamma$ such that: v is a vertex of T, T is a tree, all vertices of T are bubble vertices, and there is a unique edge conneting a vertex of T to a vertex of Γ. For example, an exterior vertex is a tree vertex. A bubble vertex which is not a tree vertex is called a **connecting vertex**. We define the **depth** of a tree vertex v to be the minimal p for which there is a sequence of tree vertices $v = v_1, \dots, v_p$ such that v_p is exterior and each v_j is connected by an edge to v_{j+1}.

A **chain of connecting vertices** is a connected saturated subgraph $R \subset \Gamma$ all of whose vertices are connecting bubbles. The name is motivated by the fact that such a graph is necessarily homeomorphic to a segment, as the reader can easily check.

We say that a bubble $C' \subset C$ is a **tree bubble** if the corresponding vertex in Γ is a tree vertex. Similarly, we define the **connecting bubbles**. Note that a tree bubble cannot contain any marked point.

Lemma 8.2. *If $C' \subset C$ is a tree bubble, then the restriction of A to C' has trivial holonomy around any node contained in C'. Hence (P, A) extends smoothly to C'.*

Proof. For the first assertion, apply induction on the depth of tree vertices, using the fact that F_A restricts to 0 on any bubble and the Chern–Weil formula in Lemma 3.5. The second statement follows from Corollary 3.2. □

Lemma 8.3. *Let $C_1, \dots, C_u$ be a sequence of connecting bubbles corresponding to a chain of connecting vertices $R \subset \Gamma$, labelled in such a way that each C_j shares a node with C_{j+1}. There exists some $\lambda \in \mathbf{i}\mathbb{R}$ with the following property. Let the exceptional points of C_j be $y_+, y_-, z_1, \dots, z_s$, where $y_\pm$ is the node which C_j has in common with $C_{j\pm1}$, except that if $j = 1$ then y_- is either a marked point or a node which C_1 shares with a principal component or another connecting bubble, or if $j = s$ then y_+ may be either a marked point or a node (connecting C_s to a principal component or a connecting bubble). Then: each z_t is a node which C_j shares with a tree bubble, the holonomy of A around any z_k is trivial, and the holonomy of A around $y_\pm$ is $e^{\pm 2\pi\lambda}$.*

Proof. Let $R \subset \Gamma$ be a maximal chain of connecting vertices. Any vertex of Γ which is connected by an edge to an interior vertex of R is necessarily a tree bubble. Furthermore, each of the two vertices at the extremes of R are connected either to a marked point vertex or to a principal vertex, and all other vertices connected to them and not contained in R are tree vertices. This explains why the exceptional points of C_j can be labelled as $y_+, y_-, z_1, \dots, z_s$ (s depends on j) and have the properties claimed in the statement of the lemma. By Lemma 8.2 the holonomy of A around z_t is trivial. So A can have nontrivial holonomy only around $y_\pm$. Finally, taking any λ such that for C_1 the holonomy around y_- is $e^{-2\pi\lambda}$, we deduce that the holonomy around the point $y_\pm \in C_j$ is $e^{\pm 2\pi\lambda}$ (use induction on j). □

8.2. **Twisted bubbles.** Let $\mathbf{y} \subset S^2$ be a finite subset. We call a **twisted bubble** over $(S^2, \mathbf{y})$ a triple (P, A, ϕ) consisting of a principal bundle $P \to S^2 \setminus \{\mathbf{y}\}$, a flat meromorphic

connection A on P and a section ϕ of $P \times_{S^1} X$ satisfying $\overline{\partial}_A \phi = 0$. We say that a twisted bubble (P, A, ϕ) is **trivial** if the covariant derivative $d_A\phi$ is identically zero.

Let $\epsilon > 0$ be a small number. Let $\Lambda^\epsilon_{\mathrm{cr}} \subset \mathbf{i}\mathbb{R}$ be the set of residues at distance $< \epsilon$ from the set or critical residues Λ_{cr}. Suppose that ϵ is small enough so that for any $\lambda \in \Lambda^\epsilon_{\mathrm{cr}}$ there is a unique critical residue $\mathrm{cr}(\lambda) \in \Lambda_{\mathrm{cr}}$ lying in the same connected component as λ. In this case, define

$$\Lambda^+_{\mathrm{cr}} = \{\lambda \in \Lambda^\epsilon_{\mathrm{cr}} \mid -\mathbf{i}\lambda > -\mathbf{i}\,\mathrm{cr}(\lambda)\}, \qquad \Lambda^-_{\mathrm{cr}} = \{\lambda \in \Lambda^\epsilon_{\mathrm{cr}} \mid -\mathbf{i}\lambda < -\mathbf{i}\,\mathrm{cr}(\lambda)\}.$$

We then have a partition $\Lambda^\epsilon_{\mathrm{cr}} = \Lambda_{\mathrm{cr}} \cup \Lambda^+_{\mathrm{cr}} \cup \Lambda^-_{\mathrm{cr}}$. Define the set of generic residues to be

$$\Lambda_{\mathrm{ge}} := \mathbf{i}\mathbb{R} \setminus \Lambda^\epsilon_{\mathrm{cr}}.$$

Theorem 8.4. *Let $\mathbf{y} = \{y_+, y_-\}$ consist of two points. There exists some $\epsilon > 0$ with the following property. Suppose that (P, A, ϕ) is a twisted bubble on $(S^2, \mathbf{y})$ satisfying $\|d_A\phi\|_{L^2} < \epsilon$. Take some local trivialization of P around y_+ and let λ be the corresponding residue of A at y_+. Then, if ϵ is small enough, we have*

1. *if $\lambda \in \Lambda_{\mathrm{cr}}$ then (P, A, ϕ) is trivial;*
2. *otherwise, by Theorem 5.1 the following limits exist*

$$\phi(y_+) := \lim_{z \to y_+} \phi(z) \in F, \qquad \phi(y_-) := \lim_{z \to y_-} \phi(z) \in F;$$

let $F_\pm$ be the connected component of F containing $\phi(y_\pm)$. Then we have
   - (a) *if $\lambda \in \Lambda^+_{\mathrm{cr}}$ then $H(F_+) \le H(F_-)$, with equality only if (P, A, ϕ) is trivial,*
   - (b) *if $\lambda \in \Lambda^-_{\mathrm{cr}}$ then $H(F_+) \ge H(F_-)$, with equality only if (P, A, ϕ) is trivial,*
   - (c) *if $\lambda \in \Lambda_{\mathrm{ge}}$ then (P, A, ϕ) is trivial.*

Proof. Suppose first that $\lambda \in \Lambda_{\mathrm{cr}}$, and write $\lambda = \mathbf{i}p/q$ for relatively prime integers p, q. Take a covering $\pi : S^2 \to S^2$ of degree q ramified at y_+ and y_-, in such a way that π^*A has trivial holonomy around y_+ and y_-. It follows from Corollary 3.2 that both P and A extend to give a bundle $P' \to S^2$ with a smooth connection A'. Then A' is flat, so trivializing any fibre of P' we obtain a trivialization of the whole bundle $P' \simeq S^2 \times S^1$, with respect to which the connection A' is trivial. This induces in particular a trivialization of P, so we can view the section ϕ as a map $\phi : S^2 \setminus \mathbf{y} \to X$, and as such it is I-holomorphic and has bounded energy. It follows from Gromov's removal of singularities theorem that ϕ extends to give a map $\Phi : S^2 \to X$. On the other hand, $\|d\Phi\|_{L^2} < \epsilon$, and if ϵ is small enough this implies that Φ is constant (this is a standard result in Gromov–Witten theory). Hence $d_A\phi = 0$, so (P, A, ϕ) is trivial.

For the remaining cases, fix a conformal isomorphism between $S^2 \setminus \mathbf{y}$ and the cylinder $\mathbb{R} \times S^1$, and take in the latter the standard coordinates (t, θ), in such a way that as t goes to $\pm\infty$ we approach the point $y_\pm$.

Consider first the case $\lambda \in \Lambda^+_{\mathrm{cr}}$. Suppose that ϵ is very small, and let (P, A, ϕ) be a bubble satisfying: $\|d_{A_u}\phi_u\|_{L^2} < \epsilon$, $d_A = d + \lambda d\theta$ in some trivialisation of P, and $-\mathbf{i}(\lambda - \lambda_{\mathrm{cr}}) \in (0, \epsilon)$. Let $\alpha := \lambda d\theta$. In the rest of the proof we are going to use the notation introduced in Section 10. Using (2) in Theorem 10.5, we can assume that both $\|d_\alpha\phi\|_{L^\infty}$ and $\|\alpha - \lambda_{\mathrm{cr}} d\theta\|_{L^\infty}$ are less than the ϵ in Theorem 11.1. By Theorem 11.1

there is some $\psi : \mathbb{R} \to X^{\lambda_{\mathrm{cr}}}$ and some $\phi_0 : \mathbb{R} \times S^1 \to TX$ such that $\phi = \exp_\psi \phi_0$. Now, Theorem 11.3 together with $d\alpha = 0$ implies that $|\psi' + \mathbf{i}\lambda I\mathcal{X}(\psi)| < Ke^{\sigma(|t|-N)}$ for every N. Making $N \to \infty$ we deduce that $\psi' = -\mathbf{i}\lambda I\mathcal{X}(\psi)$, so ψ follows a downward gradient line of H. Consequently, either $H(F_+) < H(F_-)$ or $H(F_+) = H(F_-)$, and in the latter case ψ is constantly. Now, using (11.61) in Theorem 11.1 and taking N bigger and bigger as above we deduce that $\phi_0 = 0$. It follows that, if $H(F_+) = H(F_-)$, then ϕ is constant.

The case $\lambda \in \Lambda_{\mathrm{cr}}^-$ is proved in the same way. Finally, if $\lambda \in \Lambda_{\mathrm{ge}}$ and ϵ is small enough, then we can use (10.34) in Theorem 10.3 to deduce that the image of ϕ is contained in a small and S^1-invariant ball B centered at $\phi(y_+) \in F$. Similarly, for any $\tau \in T(P, \mathbf{y})$ we have $\Phi^{\tau,A}(S^2) \subset P \times_{S^1} B$. The equivariant cohomology of B comes entirely from the classifying space. Picking τ such that $\deg P(\tau) = 0$ we deduce that $(\Phi^{\tau,A})^*[\omega(A)] = 0$. This choice of τ ensures that $\mathrm{Res}(A, y_+, \tau) + \mathrm{Res}(A, y_-, \tau) = 0$. Also, $\phi(y_-)$ belongs to the same connected component of F as $\phi(y_+)$, so $\mu(\phi(y_+)) = \mu(\phi(y_-))$. Hence, applying Lemma 7.2 to (P, A, ϕ) we deduce that $\|d_A\phi\|_{L^2} = 0$, so the bubble is trivial. □

8.3. **Proof of Theorem 8.1.** Let $\mathcal{C} = ((C, \mathbf{x}), (P, A, G), \phi, \{\mathcal{T}_y\}, \{\mathcal{T}_x\})$ be a c-STHM of genus g and n marked points satisfying $\mathcal{YMH}_c(\mathcal{C}) \leq K$. A bubble $C' \subset C$ is said to be **unstable** if it contains less than 3 exceptional points. Let ϵ be as in Theorem 8.4. (Note that a particular consequence is that if $\psi : S^2 \to X$ is a I-holomorphic map with energy less than ϵ, then ψ is trivial — this corresponds to the case of trivial holonomy around $y_\pm$ in the lemma.)

Let Γ be the graph associated to $(C, \mathbf{x})$. All subgraphs of Γ which we shall mention will be saturated. The subgraph of Γ consisting of connecting vertices can be written as the disjoint union of the set of maximal chains of connecting vertices $R_1, \dots, R_l$. Furthermore, l is at most equal to the number of marked points $\mathbf{x}$ plus the number of nodes of C (because the bubbles corresponding to vertices in R_j can be identified with the connecting bubbles which are contracted by the stabilisation map either to a given marked point or to a node). Hence, $l \leq 3g - 3 + n$.

Each tree vertex of Γ belongs to a maximal tree $T \subset \Gamma$, all of whose vertices are tree vertices. Hence, the subgraph of Γ consisting of tree vertices is the union of trees $T_1, \dots, T_r$. Each such tree T_j has a distinguished vertex, which we call the **root**, which is connected by an edge to a vertex of $\Gamma \setminus T_j$. If T is a tree, we say that a vertex of T is stable (resp. unstable) if its degree is ≥ 3 (resp. ≤ 2). Let $|T|$ denote the number of vertices of T. Then the number of unstable vertices of T is at least $(|T| + 2)/3$. This follows from estimating the number of stable vertices by counting edges: T has $|T| - 1$ edges and each stable vertex contributes at least $3/2$ to the total number of edges. Now, an unstable vertex of a tree T_j corresponds to an unstable bubble of C, unless the vertex is the root of T_j and it has degree 2 (hence its degree as a vertex of Γ is 3). In this case, the root can not be the unique unstable vertex of T_j, so there is at least one unstable bubble for each tree T_j. Since each such bubble is nontrivial, it contributes at least ϵ to the total energy $\mathcal{YMH}_c(\mathcal{C})$. Hence, the number of trees can be bounded as $r \leq \epsilon^{-1}K$. On the other hand, the previous arguments tell us that for each T_j there are at least $(|T_j| - 1)/3$ unstable bubbles (we are substracting here a unit in case the root does not

correspond to an unstable bubble). Hence we have $\sum_{j=1}^{r}(|T_j|-1)/3 \leq \epsilon^{-1}K$, which combined with the bound on r yields $\sum_{j=1}^{r}|T_j| \leq \epsilon^{-1}4K$. Hence, the number of tree bubbles in C is bounded.

It only remains to prove that the length of any chain of connecting vertices of Γ is bounded by a constant independent of $\mathcal{C}$. Let $R \subset \Gamma$ be any such chain. A vertex of R will be said to be stable if its degree *as a vertex of* Γ is ≥ 3, and unstable otherwise. Each interior vertex of R which is stable is connected by an edge to at least one of the trees T_j. Since each tree T_j is connected by an edge to a unique vertex in $\Gamma \setminus T_j$, it follows that the number of interior stable vertices in R is at most $\epsilon^{-1}K$. So if we prove that each sequence of consecutive unstable vertices of R is bounded, it will follow that the number of vertices of R is bounded. Now, this follows from combining Lemma 8.3 with Theorem 8.4, and the fact that the number of connected components of X is finite.

9. Main theorem: compactness

The following is the main result of the paper.

Theorem 9.1. *Let g and n be nonnegative integers satisfying $2g+n \geq 3$. Let $K>0$ be any number, and let $c \in \mathbf{i}\mathbb{R}$. Let $\{\mathcal{C}_u\}$ be a sequence of c-stable twisted holomorphic maps of genus g and with n marked points, satisfying $\mathcal{YMH}_c(\mathcal{C}_u) \leq K$ for each u. Then there is a subsequence $\{[\mathcal{C}_{u_j}]\}$ converging to the isomorphism class of another c-stable twisted holomorphic map $\mathcal{C}$. Furthermore, we have*

$$\lim_{j\to\infty} \mathcal{YMH}_c(\mathcal{C}_{u_j}) = \mathcal{YMH}_c(\mathcal{C}). \tag{9.24}$$

The rest of this section is devoted to the proof of the theorem. Many ideas involved in the proof are the same that appear in the compactness theorem for stable maps in Gromov–Witten theory, which we will assume that the reader is familiar with (see for example [T, RT1, IS, FO]). First of all, note that if a subsequence satisfies $\mathcal{C}_{u_j} \to \mathcal{C}$ then, by Theorem 7.4, (9.24) holds automatically.

9.1. Getting the first limiting curve.

Let $(C_u, \mathbf{x}_u)$ be the nodal marked curve underlying $\mathcal{C}_u$, so that $\mathbf{x}_u$ is the union of the critical points $\mathbf{x}_{\mathrm{cr},u}$ and the generic ones $\mathbf{x}_{\mathrm{ge},u}$. By Theorem 8.1 the number of bubbles in each C_u is uniformly bounded. It follows that we can assume (passing to a subsequence) that all curves C_u have the same topological type. For any u, let $\mathbf{x}_u^0 \subset C_u$ be a list of points such that, setting $\mathbf{x}'_u := \mathbf{x}_u \cup \mathbf{x}_u^0$, the marked curve $(C_u, \mathbf{x}'_u)$ is stable, and suppose that each $\mathbf{x}_u^0$ has as few elements as possible. Taking a subsequence, we can assume that $(C_u, \mathbf{x}'_u)$ converges to a stable curve $(C', \mathbf{x}')$. The new set of marked points $\mathbf{x}'$ contains the limit $\mathbf{x}$ of the sequences $\mathbf{x}_u \subset \mathbf{x}'_u$. We call the points in $\mathbf{x}$ **original marked points**. It is necessary to make this distinction because the connection may have poles in the original marked points, whereas in the other marked points it will always be smooth. We remark also, to avoid confusion, that in the course of the proof the lists of marked points $\mathbf{x}'_u$ and $\mathbf{x}'$ will increase, and that the curve C' will change from time to time (the changes will be addition of rational components).

In the names of the following three sections (where we describe how bubbles are to be added to the limit curve C') we use the terminology of Section 8.1.

9.2. **Adding tree bubbles, first part.** We first consider bubbling off away from nodes and marked points. Following the approach in [FO], we force the appearence of bubbles in the domains $(C_u, \mathbf{x}'_u)$ by adding new marked points near points where $|d_{A_u}\phi_u|$ blows up. (In order to have a choice of Riemannian metric on each stable curve — which we need to define $|d_{A_u}\phi_u|$ — we chose an element in $\mathcal{M}et_{g,n'}$, where n' is the number of points in $\mathbf{x}'$, see Section 5.4. Later on we will use metrics on stable curves with more than n' marked points, and it will be implicitly assumed that we have chosen an element of the corresponding space $\mathcal{M}et$.) Suppose that $K \subset C'$ is a compact set which does not contain any original marked point nor any node. Since $(C_u, \mathbf{x}'_u)$ converges to $(C', \mathbf{x}')$, we can assume, provided u is big enough, that there is a canonical inclusion of $\iota_K : K \to C_u$, whose image we call K_u (see Section 6.3). Assume that $\sup_{K_u} |d_{A_u}\phi_u|$ is not bounded as u goes to infinity. Then we pick for each u a point $x_u^1 \in K_u$ where $|d_{A_u}\phi_u|$ attains its supremum, which we denote by s_u, and another point x_u^2 at distance s_u^{-1} from x_u^1. We add x_u^1 and x_u^2 to the list x'_u and obtain a new marked nodal curve still denoted by $(C_u, \mathbf{x}'_u)$. Passing to a subsequence, we can assume that $\{(C_u, \mathbf{x}'_u)\}$ converges to another limit curve $(C', \mathbf{x}')$. Then we repeat the process: we take a compact set $K \subset C'$ not containing original marked points nor smooth points, and so on. The process stops when, for any such K, the sequence $\sup_{K_u} |d_{A_u}\phi_u|$ is bounded. That the process stops after adding a finite number of points is proved exactly as in Gromov–Witten theory (see for example Proposition 11.3 in [FO]): namely, each time we add two points, there appears a new bubble, and each bubble which appears contributes more than a certain amount $\epsilon > 0$ to the energy $\mathcal{YMH}_c(\mathcal{C}_u)$, so there cannot be infinitely many. The key point here is that each A_u is smooth on K_u and has bounded curvature. Hence, when we zoom up the connections become more and more flat (this is because we are in real dimension 2), and in the limit they become trivial. So the bubbles which appear in this situation are actually holomorphic maps in the usual sense.

9.3. **Adding tree bubbles, second part.** Suppose that the limit curve $(C', \mathbf{x}')$ has k nodes, so that the curves of the form $(C_u, \mathbf{x}'_u)$ can be identified with $C'(I', \delta_{u,1}, \ldots, \delta_{u,k}, \mathbf{x}')$. Now take some j such that $\delta_u := \delta_{u,j}$ is nonzero for big enough u. We say in this case that the node is **new**, i.e., it does not appear in the curves C_u if u is big. Let $N_u := N_j(\delta_{u,j})$ be the neck defined in Section 6.3. We fix an isomorphism

$$N_u \simeq [\ln|\delta_u| - \ln \epsilon, \ln \epsilon] \times S^1 =: [p_u, q_u] \times S^1$$

and we denote as usual by t, θ the standard cylindrical coordinates. Also, we take on N_u the standard cylindrical metric. For any $\Delta > 0$, let $N_u(\Delta) := [p_u + \Delta, q_u - \Delta]$. Suppose that

$$\limsup_{\Delta\to\infty} \left(\limsup_{u\to\infty} \sup_{N_u(\Delta)} |d_{\alpha_u}\phi_u| \right) = \infty. \tag{9.25}$$

In this case, we proceed as in §9.2, picking pairs of sequences of points x_u^1 and x_u^2 at larger and larger distance from the boundary of N_u, such that $s_u := |d_{\alpha_u}\phi_u(x_u^1)|$ goes

to infinity and x_u^2 is at distance s_u^{-1} from x_u^1. We then add x_u^1 and x_u^2 to $\mathbf{x}'_u$, go to a subsequence so that there is a new limit $(C_u, \mathbf{x}'_u) \to (C, \mathbf{x}')$, and repeat the process. We repeat this, going again to the beginning of §9.3, as many times as possible, and at some point we must stop.

In case the j-th node was not new, that is, $\delta_u = 0$ for big enough u, we consider the normalization of each C_u near the node, and apply the same technique as in the previous case in a neighborhood of each of the two preimages of the node. The only difference in this case is that instead of having a finite cylinder N_u we will have a semiinfinite cylinder $N_u \simeq [\ln \epsilon, \infty) \times S^1$, and we define $N_u(\Delta) := [\ln \epsilon + \Delta, \infty) \times S^1$.

Finally, we apply the same technique around each original original marked point, modelling again a punctured neighborhood of it with a semiinfinite cylinder $[0, \infty) \times S^1$.

9.4. **Connecting bubbles appear.** Take again some node in C' and assume that it is new. Define the cylinders N_u and $N_u(\Delta)$ and in §9.3. Since we repeated the process in §9.3 as many times as possible, we must have now

$$\limsup_{\Delta \to \infty} \left(\limsup_{u \to \infty} \sup_{N_u(\Delta)} |d_{\alpha_u} \phi_u| \right) < \infty.$$

This means that for some $\Delta > 0$ we have $\limsup_{u \to \infty} \sup_{N_u(\Delta)} |d_{\alpha_u} \phi_u| < \infty$. We now replace N_u by $N_u(\Delta)$ and denote the extremes of the new cylinder N_u again by p_u and q_u.

The neck $N_u \subset C_u$ can either belong to a principal or a bubble component. Suppose we are in the first case. Then the vortex equation $\iota_{d\,\mathrm{vol}(\nu_u)} F_A + \mu(\phi) = c$ is satisfied (ν_u denotes the restriction of the metric on the componenet C_u to which N_u belongs). Let us write $d\,\mathrm{vol}(\nu_u) = f_u dt \wedge d\theta$. We have exponential bounds on each derivative of f: for any $l > 0$,

$$|\nabla^l f_u(z)| \leq K_l e^{-d(z, \partial N_u)}, \tag{9.26}$$

where K_l is independent of u. This follows from the fact that ν_u is the restriction to C_u of a smooth metric on the universal curve over a moduli space of stable curves, which is a compact orbifold; hence all its derivatives are bounded. On the other hand, the vortex equation takes the following form:

$$d\alpha_u = f_u(c - \mu(\phi_u)). \tag{9.27}$$

Let $\epsilon > 0$ be smaller than the ϵ's in Theorems 10.3 and 11.1. For any u, let $d_u : [p_u, q_u] \to \mathbb{R}$ be the function defined as

$$d_u(t) := \sup_{\{t\} \times S^1} |d_{\alpha_u} \phi_u|.$$

We define an **ϵ-bubbling list** to be a list of sequences $(\{b_u^1\}, \ldots, \{b_u^r\})$ satisfying:

(1) each b_u^j belongs to N_u,

(2) denoting $b_u^0 := p_u$ and $b_u^{r+1} := q_u$ we have, for each j between 0 and r,

$$\lim_{u\to\infty} b_u^{j+1} - b_u^j = \infty.$$

(3) for each j we have $\liminf_{u\to\infty} d_u(b_u^j) \geq \epsilon$.

Lemma 9.2. *The number of sequences in an ϵ-bubbling list is bounded in terms of $K := \sup_u \mathcal{YMH}_c(\mathcal{C}_u)$.*

Proof. For any $b \in [p_u+1/2, q_u-1/2]$, let $C(b)$ be the cylinder $[b-1/2, b+1/2]\times S^1$. The lemma follows from this claim: there exists some $\eta > 0$ (independent of $\{\mathcal{C}_u\}$) such that, whenever $d_u(b) \geq \epsilon$, we have $\|d_{\alpha_u}\phi_u\|_{L^2(C(b))} \geq \eta$. Indeed, then the length of a ϵ-bubbling list is at most $\eta^{-1}K$. To prove the claim, fix some $K_0 > 0$. If $\|d_{\alpha_u}\phi_u\|_{L^2(C(b))} \leq K_0$, then putting momentarily the connections in Coulomb gauge and using elliptic bootstrapping alternatively with equation (9.27) and $\overline{\partial}_{\alpha_u}\phi_u = 0$ (taking the bound (9.26) into account) we deduce a uniform bound on the L^2_2 bound of $d\alpha_u$. Since $L^2_2 \subset L^p_1$ for any p, we may apply Corollary 10.15 to deduce that $\|d_{\alpha_u}\phi_u\|_{L^2(C(b))} \geq \delta$ for some $\delta > 0$ independent of $\{\mathcal{C}_u\}$. Finally, we set η to be the minimum of δ and K_0. □

Take an ϵ-bubbling list $(\{b_u^1\}, \dots, \{b_u^r\})$ of maximal length. If the list is empty, then we do nothing, and begin again the process in §9.4 with another node. If there is no node near which we can construct a nonempty ϵ-bubbling list, then we pass to the next step in §9.6. If instead the list is nonempty, then we define the points $x_u^{i,1} := (b_u^i, 1) \in N_u$ and $x_u^{i,2} := (b_u^i + 1, 1) \in N_u$. We add to $\mathbf{x}'_u$ these $2r$ new points and do as always: pass again to a subsequence so that there is a new limiting curve $(C_u, \mathbf{x}'_u) \to (C', \mathbf{x}')$ and begin again with a node of the new curve C'. The process has to stop at some moment because the energy is bounded.

As in §9.3, we do the same for the nodes of C' which are not new: take the normalization and do what we did in the cylinders N_u in a neighborhood of each preimage of the node, which is conformally equivalent to a semiinfinite cylinder. Finally, we consider neighbourhoods of each original marked point and do exactly the same.

9.5. **Vanishing old bubbles.** After all this process of adding marked points at the curves C_u we end up with a limiting curve $(C', \mathbf{x}')$. We denote from now one $C := C'$. Suppose that C has k nodes (this k will most of the times be bigger than the one in §9.3), so that for big enough u there is an isomorphism of marked curves

$$\xi_u : C(I', \delta_{u,1}, \dots, \delta_{u,k}, \mathbf{x}) \to (C_u, \mathbf{x}_u).$$

We say that a bubble $C_0 \subset C$ is **old** if for each node z_j contained in C_0 the smoothing parameters $\delta_{u,j}$ vanish. This means that the bubble C_0 already existed in the curves C_u, so for any u we have an inclusion $C_0 \subset C_u$. We say that an old bubble C_0 **vanishes in the limit** if for big enough u we have $d_{A_u}\phi_u|_{C_0} \neq 0$ and

$$\lim_{u\to\infty} \|d_{A_u}\phi_u\|_{L^2(C_0)} = 0.$$

9.6. **Constructing the limiting c-STHM.** Take a compact set $K \subset C$ disjoint from the nodes of C and the original marked points. Since we have the bound

$$\sup_u \sup_K |d_{\xi_u^* A_u} \xi_u^* \phi_u| < \infty,$$

standard arguments (for example, Lemma 10.13 combined with a patching argument as in §4.4.2 in [DK]) imply that there is a subsequence of the sequence $\xi_u^*(P_u, A_u, \phi_u)$ which, after regauging, converges to a limiting triple (P_K, A_K, ϕ_K). Taking an exhaustion of the smooth locus of $C \setminus \mathbf{x}$ by compact sets, we obtain a limiting triple (P, A, ϕ) which satisfies the vortex equations on principal components, and such that ϕ is holomorphic with respect to A. Furthermore, $\|d_A \phi\|_{L^2}$ is finite.

The triple (P, A, ϕ) will be part of the limiting c-STHM $\mathcal{C}$. Let $\mathbf{z}$ be the set of nodes of C. By construction the bundle P is defined over $C \setminus (\mathbf{x} \cup \mathbf{z})$, and we now prove that A is meromorphic. On each principal component the pair (A, ϕ) satisfies the equations

$$\overline{\partial}_A \phi = 0 \qquad \text{and} \qquad \iota_{d\,\mathrm{vol}(\nu_{[C^{\mathrm{st}},\mathbf{x}]})} F_A + \mu(\phi) = c \tag{9.28}$$

(here C^{st} is the stabilization of C). The second equation implies that the curvature of A is bounded, and the first one, combined with the fact that the energy $\|d_A \phi\|_{L^2}$ is finite, allows to apply Corollary 10.2 and deduce that $\mu(\phi)$ extends continuously to C. Going back now to the second equation again, we deduce that F_A extends continuously to C, so A is meromorphic. On the other hand, the restriction of F_A to each of the bubbles of C is zero. Finally, the stability condition in the bubbles is satisfied.

Hence, it only remains to construct the limiting gluing data G for P and the collections of chains of gradient segments $\{\mathcal{T}_y\}$, $\{\mathcal{T}_x\}$. This will be done in the next three sections.

9.7. **Chains of gradient segments and gluing data at the old nodes.** Let y, y' be the preimages of an old node in the normalization of C. Then y and y' also belong to each of the normalizations of the curves C_u, hence we have chains of segments of gradient lines $\mathcal{T}_{u,y}$ and $\mathcal{T}_{u,y'}$. Since the space of gradients of segment lines is compact, we can assume that there are limit chains $\mathcal{T}_y$ and $\mathcal{T}_{y'}$. The same happens with gluing data.

On the other hand, each vanishing old bubble $C_0 \subset C$ gives rise to an infinite gradient segment $\mathcal{T}$ in X by Theorem 8.4. If C_0 has only two exceptional points (it can't have only one, because then it would be a tree bubble), then we collapse C_0. This identifies two different nodes in C. Suppose that $\mathcal{T}_a$ and $\mathcal{T}_b$ are the chains of gradient segments in each of them (taken with the same orientation). Then the chain of segments in the new node is the concatenation of $\mathcal{T}_a$, $\mathcal{T}$ and $\mathcal{T}_b$. The gluing data is the obvious one. We leave the details of this construction to the reader.

If C_0 has more than two exceptional points, $y_+, y_-, z_1, \ldots, z_r$, and the poles of A_u are in $y_\pm$, then we get as before a limit gradient segment $\mathcal{T}$. But instead of collapsing C_0 we substitute it by a chain of r trivial bubbles, each of them containing one of the points $z_1, \ldots, z_r$ and two nodes. And in the node shared by the bubbles containing z_j and z_{j+1} we take the portion of $\mathcal{T}$ which lies between the limits $\lim_{u\to\infty} S^1 \cdot \phi(z_j)$ and $\lim_{u\to\infty} S^1 \cdot \phi(z_{j+1})$. Again we leave the details to the reader.

9.8. **Chains of gradient segments and gluing data at the new nodes.** Take some new node $z \in \mathbf{z}$ with preimages y, y' in the normalisation of C. For each u let, as always, $N_u := N_y(\delta_{u,z}) \subset C(I', \{\delta_{u,z}\}, \mathbf{x}') \simeq C_u$ be the neck stretching to z. Assume that $N_u = [p_u, q_u] \times S^1$ and define $N_u(\Delta) := [p_u + \Delta, q_u - \Delta] \times S^1$. There is some $\Delta_0 > 0$ such that for any $\Delta \geq \Delta_0$ we have

$$\limsup_{u \to \infty} \sup_{N_u(\Delta)} |d_{A_u} \phi_u| < \epsilon, \tag{9.29}$$

where $\epsilon > 0$ is less than the ϵ's in Theorems 10.3 and 11.1.

Assume that y is in the side of $\{p_u\} \times S^1$ and y' in the side of $\{q_u\} \times S^1$. By Corollary 10.2 there is a limiting triple (P_y, A_y, ϕ_y) defined over S_y, and similarly for y'. Let $\mathcal{O} \subset X$ be the orbit on which ϕ_y takes values, and define $\mathcal{O}'$ similarly. It follows from our construction that

$$\lim_{\Delta \to \infty} \left(\limsup_{u \to \infty} d_{S^1}(\phi(\{p_u + \Delta\} \times S^1), \mathcal{O}) \right) = 0, \tag{9.30}$$

and similarly for $\mathcal{O}'$.

Pick a trivialization of P_u on N_u with respect to which $d_{A_u} = d + \alpha_u$ is in balanced temporal gauge with residue λ_u (see Section 6.4). We can assume that $|\lambda_u| \leq 1$ and, passing to a subsequence, that there is a limit $\lambda_u \to \lambda$. We distinguish two possibilities.

(1) Suppose first that λ is not critical, so that both $\mathcal{O}$ and $\mathcal{O}'$ lie in the fixed point set. In this case we can take the chains $\mathcal{T}_y$ and $\mathcal{T}_{y'}$ to be degenerate, and in fact we have $\mathcal{O} = \mathcal{O}'$. To see this it suffices to prove that

$$\lim_{\Delta \to \infty} \left(\limsup_{u \to \infty} \operatorname{diam}(\phi_u(N_u(\Delta))) \right) = 0.$$

Now, this formula follows from applying Theorem 10.3. Indeed, on the one hand we have for any $z = (t, \theta) \in N_u(\Delta_0)$ a bound

$$|d_\alpha \phi(z)| \leq K e^{-\sigma d(z, \partial N_u(\Delta_0))} \epsilon.$$

Now, if $\Delta \geq \Delta_0$ and $z \in N_u(\Delta)$ we have, for the same reason,

$$|d_\alpha \phi(z)| \leq K e^{-\sigma(\Delta - \Delta_0)} e^{-\sigma d(z, \partial N_u(\Delta))} \epsilon.$$

In particular, $\|d_\alpha \phi\|_{L^\infty(N_u(\Delta))}$ goes to 0 as $\Delta \to \infty$. Combined with (10.34), we deduce that

$$\lim_{u \to \infty} \operatorname{diam}_{S^1}(\phi_u(N_u(\Delta))) = 0$$

which, using (9.30) and the fact that $\mathcal{O}$ is a fixed point, implies that the actual diameter goes to zero:

$$\lim_{u \to \infty} \operatorname{diam}(\phi_u(N_u(\Delta))) = 0$$

(2) Now suppose that λ is critical. We have to construct for big enough u a cylinder $M_u := N_u(\Delta_u) \times S^1$ such that the conditions given in Section 6.5 (subsection

Convergence near the nodes to chains of gradient segments) are satisfied. Observe first of all that for any $\Delta > 0$ there is some $u(\Delta)$ such that if $u \geq u(\Delta)$ then

$$M_u^\Delta := ([p_u, p_u + \Delta] \cup [q_u - \Delta, q_u]) \times S^1$$

is contained in the smooth locus of the limit curve C. In particular, ϕ restricts to give a map $\phi : M_u^\Delta \to X$. On the other hand, for any $\epsilon > 0$ there is some $u(\Delta, \epsilon) \geq u(\Delta)$ such that if $u \geq u(\Delta, \epsilon)$ then

$$\sup_{z \in M_u^\Delta} d_{S^1}(\phi(z), \phi_u(z)) < \epsilon$$

(because ϕ_u converge modulo gauge to ϕ). Now take sequences $\Delta_r \to \infty$ and $\epsilon_r \to 0$, and define for every r the number $u_r := u(\Delta_r, \epsilon_r)$. We can assume without loss of generality that the sequence $\{u_r\}$ is strictly increasing. Now, take any $u > 0$ and define r by the condition $u_r \leq u < u_{r+1}$. Then define $\Delta_u := \Delta_r$ and also $M_u := N_u(\Delta_u)$. Given Δ and a big enough u, let $M_u^{\Delta,\pm}$ be the two connected components of $N_u(\Delta) \setminus M_u$. It follows from the construction that

$$\lim_{\Delta \to \infty} \left(\limsup_{u \to \infty} \operatorname{diam}_{S^1}(\phi_u(M_u^{\Delta,\pm})) \right) = 0.$$

So it remains to prove that the sequence of cylinders (ϕ_u, M_u) converges with gauge λ to a certain chain $\mathcal{T} \in \mathcal{T}(X^\lambda)$.

If u is big enough and $\Delta \geq \Delta_0$, Theorem 11.1 gives us a map $\psi_u(\Delta) : T_u(\Delta) := [p_u + \Delta, q_u - \Delta] \to X$. Let $l_u := -\mathbf{i}(\lambda_u - \lambda)$. Formula (11.63) in Theorem 11.3 (applied first to $T_u(\Delta_0)$ and then to $T_u(\Delta)$) implies that

$$|\psi_u(\Delta)'(t) - l_u I(\psi(t)) \mathcal{X}(\psi_u(\Delta)(t))| < K e^{-\sigma(\Delta - \Delta_0)} e^{-\sigma d(t, \partial S_u(\Delta))} K \tag{9.31}$$

for some constant K. Let $M_u = T_u \times S^1$. Since $\Delta_u \to \infty$, the previous inequality allows us to apply Theorem 12.1. Hence, passing to a subsequence, we can distinguish two possibilities.

- If $l_u|T_u| \to 0$ then $\operatorname{diam}(\psi_u(T_u)) \to 0$, and, passing again to a subsequence, we may define $\mathcal{T}_y$ to be the degenerate chain of gradient segments with constant value the unique point $x \in X$ such that $\lim d(\psi_u(T_u), x) = 0$.
- If $\lim l_u|T_u| \neq 0$ then for big enough u we have $l_u \neq 0$, and by the theorem $(l_u^{-1}\psi_u, l_u T_u)$ converges to a chain of gradient segments, which we denote by $\mathcal{T}_y$.

Inequality (11.60) and formula (11.58) in Theorem 11.1 imply that the renormalized sequence $(l_u^{-1}\phi_u, l_u M_u)$ converges with gauge λ to $\mathcal{T}_y$.

Finally, to define the gluing data one only needs to pass to a suitable subsequence, as is clear from the conditions *Convergence of gluing angles* and *Convergence of gluing data* in the definition of convergence of sequences of c-STHM (Section 6.5). Indeed, the set of possible gluing data in a c-STHM is compact.

9.9. **Chains of gradient segments in generic marked points.** Let $\{x_u\}$ be a sequence of original marked points, where each $x_u \in \mathbf{x}_{\mathrm{ge},u}$. Let $\lambda_u \in \mathbf{i}\mathbb{R}$ satisfy $e^{2\pi\lambda} = \mathrm{Hol}(A_u, x_u)$, and let $\mathcal{T}_u$ be the chain of gradient segments $\mathcal{T}_{x_u}$. Suppose that $\{x_u\}$ converges to $x \in C'$. This means that there is a compact subset $K \subset C \setminus (\mathbf{x} \cup \mathbf{z})$ such that for big enough u we have $x_u \subset \iota_K(K) \subset C_u$. Passing to a subsequence we can also assume that $\lambda_u \to \lambda \in \mathbf{i}\mathbb{R}$. In this section we show that, passing to a subsequence, there is a well defined limiting chain of gradient segments $\mathcal{T}_x$. The analysis is very similar to the previous one, so we will be sketchy. We distinguish two situations.

(1) Suppose that λ is not critical. In this case we define $\mathcal{T}_x$ to be the unique degenerate chain which satisfies the matching condition with ϕ_x. (Note that for big enough u the residue λ_u is also noncritical, hence $\mathcal{T}_u$ is degenerate.)
(2) Suppose that λ is critical. In this case, passing to a subsequence, we can assume that one of the following possibilities holds:
   - either λ_u is critical for big enough u, in which this case we define $\mathcal{T}_x$ to be the limit of the sequence of chains $\{\mathcal{T}_u\}$,
   - or λ_u is not critical for big enough u, and we follow the same strategy as in case (2) of the previous section to define $\mathcal{T}_x$.

10. Local estimates and cylinders with noncritical residue

Theorem 10.1. *Let P be a principal S^1 bundle on the punctured disk $\mathbb{D}^*$, and fix an element $\tau \in T(P, 0)$. Assume that A is a connection on P whose curvature is uniformly bounded, $|F_A|_{L^\infty} < \infty$, and such that $\mathrm{Res}(A, 0, \tau) \notin \mathbf{i}\mathbb{Z}$. Let ϕ be a section of the trivial bundle $P \times_{S^1} X$ which satisfies $\overline{\partial}_A\phi = 0$ and $\|d_A\phi\|_{L^2(\mathbb{D}^*)} < \infty$. Then, denoting by (r, θ) the polar coordinates on $\mathbb{D}^*$, there are constants $K > 0$ and $\nu > 0$ such that*

$$|d_A\phi(r, \theta)| \leq Kr^\nu.$$

The proof of Theorem 10.1 will be given in Section 10.4.

Corollary 10.2. *Under the hypothesis of Theorem 10.1, the section ϕ extends at the origin to give a section ϕ_0 of Y_0. In particular, the composition of ϕ with the moment map, $\mu(\phi)$, extends to a continuous map from the disk $\mathbb{D}$ to $\mathbf{i}\mathbb{R}$. Furthermore we have:*

(1) *The section ϕ_0 is covariantly constant with respect to the limiting connection A_0 on P_0, that is, $d_{A_0}\phi_0 = 0$.*
(2) *Let $\lambda := \mathrm{Res}(A, 0, \tau)$. The section ϕ_0 takes values in X^λ. In particular, if λ is not critical, then ϕ_0 takes values in the fixed point set F, hence ϕ_0 is constant and the following limit exists*

$$\phi(0) := \lim_{z \to 0} \phi(z) \in F. \tag{10.32}$$

Proof. We prove (1). Take a trivialisation of P around 0 for which $d_A = d + \alpha + \lambda d\theta$, where α is of type C^1 on the whole disk $\mathbb{D}$ and is in radial gauge, i.e., $\alpha = ad\theta$ for some function a on $\mathbb{D}$ vanishing at the origin. The trivialisation of P induces a trivialisation of $P_0 \to S_0$, with respect to which $d_{A_0} = d + \lambda d\theta$. Using the trivialisation of P we

look at the section ϕ as a map $\phi : \mathbb{D}^* \to X$. Now define for every $0 < r < 1$ the map $\phi_r : S^1 \to X$ by $\phi_r(\theta) := \phi(r, \theta)$. It follows from the estimate Theorem 10.1 that the limit $\phi_0 := \lim_{r\to 0} \phi_r$ exists and is of type C^1. Define now the connection A_r on the trivial S^1 bundle over the circle using the 1-form $\alpha_r(r, \theta) := ra(r, \theta)d\theta$. It follows also from Theorem 10.1 that $d_{A_0}\phi_0 = \lim_{r\to 0} d_{A_r}\phi_r = 0$. Finally, (2) follows from (1), observing that ϕ_0 takes values in a unique orbit of the action of S^1 on X. □

To study the local properties of equation (2.8) we restrict ourselves to considering trivial principal S^1 bundles P over a (nonnecessarily compact C). Then $Y = C \times X$, the sections of Y can be identified with maps $C \to X$, and the connections on Y are the same as forms $\alpha \in \Omega^1(C, \mathbf{i}\mathbb{R})$. Finally, there is a canonical splitting $TY = TC \oplus TX$ (we omit the pullbacks). With respect to a given form α the bundle of horizontal tangent vectors is $T^{\text{hor}}_\alpha = \{(u, \mathbf{i}\alpha(u)\mathcal{X}) \mid u \in TC\} \subset TY$. Consequently, the covariant derivative of a section $\phi : C \to X$ is $d_\alpha\phi = d\phi - \mathbf{i}\alpha\mathcal{X}(\phi)$. Finally, if I is an almost complex structure on X, we have

$$\overline{\partial}_{I,\alpha}\phi = \overline{\partial}_I\phi - \frac{1}{2}(\mathbf{i}\alpha\mathcal{X}(\phi) + \mathbf{i}(\alpha \circ I_C)(I\mathcal{X})(\phi)).$$

A pair (α, ϕ) consisting of a 1-form α on C with values on $\mathbf{i}\mathbb{R}$ and a map $\phi : C \to X$ will be simply called a **pair**; to specify both the curve C and the target manifold X we will write

$$(\alpha, \phi) : C \to X.$$

If I is any S^1-invariant almost complex structure on X, we will say that the pair (α, ϕ) is **I-holomorphic** if the equation $\overline{\partial}_{I,\alpha}\phi = 0$ is satisfied.

For any natural number N, denote $C_N := [-N, N] \times S^1$. We can now state the second main result of this section.

Theorem 10.3. *For any noncritical $\lambda \in \mathbf{i}\mathbb{R} \setminus \Lambda_{\text{cr}}$ there exist some $K > 0$, $\sigma > 0$ and $\epsilon > 0$, depending continuously on λ, with the following property. Let $(\alpha, \phi) : C_N \to X$ be a I-holomorphic pair satisfying $\|\alpha - \lambda d\theta\|_{L^\infty} < \epsilon$ and $\|d_\alpha\phi\|_{L^\infty(C_N)} < \epsilon$. Then the following inequality holds for any t, θ:*

$$|d_\alpha\phi(t, \theta)| \leq Ke^{-\sigma(N-|t|)}\|d_\alpha\phi\|_{L^\infty(C_N)}. \tag{10.33}$$

In particular this implies

$$\operatorname{diam}_{S^1}(\phi(C_N)) < K\|d_\alpha\phi\|_{L^\infty(C_N)}. \tag{10.34}$$

The proof of Theorem 10.3 will be given in Section 10.5

10.1. **Equivariant charts.** Let I_0 (resp. g_0) denote the standard complex structure on $\mathbb{C}^n$, viewed as a differentiable manifold.

Lemma 10.4. *There exists a finite set $W = W(X) \subset \mathbb{Z}^n$ and a real number $r = r(X, \omega, I) > 0$ with the following property. For any $\epsilon > 0$ and any $x \in X$ there exists an equivariant open neighbourhood U of x, an action of S^1 on $\mathbb{C}^n$ whose collection of weights belongs to $W(X)$, an equivariant almost complex structure I_x and metric g_x on $\mathbb{C}^n$ satisfying $\|I_x - I_0\|_{L^\infty} < \epsilon$ and $\|g_x - g_0\|_{L^\infty} < \epsilon$ (both norms taken with respect to*

g_0), and an equivariant map $\xi : U \to \mathbb{C}^n$ which is a complex isometry. Furthermore, we can assume that $\xi(U)$ is contained in the ball $B(0, 2r) \subset \mathbb{C}^n$.

Proof. Take r to be the length of the longest orbit in X divided by 2π. If x is a fixed point the chart is easily constructed using the exponential map with respect to the invariant metric g. Suppose that x is not a fixed point, and that the length of the orbit through x is $2\pi\rho$. Let Γ be the stabiliser of x, which acts linearly on T_xX. Let $L \subset T_xX$ be the complex subspace generated by $\mathcal{X}(x)$ and $I\mathcal{X}(x)$ (recall that $\mathcal{X} \in \Gamma(TX)$ is the vector field generated by the infinitesimal action of $\operatorname{Lie} S^1$). Both $\mathcal{X}(x)$ and $I\mathcal{X}(x)$ are fixed by the action of Γ, so L is Γ-invariant. Let $N \subset T_xX$ be the Hermitian ortogonal of L, which is a Γ-invariant and complex vector subspace. Take a complex isomorphism $\mathbb{C}^{n-1} \simeq N$. Consider the induced action of Γ on $\mathbb{C}^{n-1}$. Fix any linear extension ρ of this action to S^1, which we may assume (chosing appropiatedly the identification $\mathbb{C}^{n-1} \simeq N$) to be diagonal and compatible with the standard metric in $\mathbb{C}^{n-1}$. Let $k \in \mathbb{N}$ be the order of Γ, consider the action of S^1 on $\mathbb{C}^* \times \mathbb{C}^{n-1}$ given by $\theta \cdot (x, y) := (\theta^k x, \rho(\theta)(y))$. There exists an S^1-invariant neighbourhood $R \subset \mathbb{C}^*$ of the circle S_ρ of radius ρ centered at the origin in $\mathbb{C}$ and an S^1-equivariant embedding

$$f : R \times \mathbb{C}^{n-1} \to X$$

(here we view $R \times \mathbb{C}^{n-1}$ as a subset of $\mathbb{C}^* \times \mathbb{C}^{n-1}$) which satisfies, for any $t \in \mathbb{R}$ and $n \in \mathbb{C}^{n-1}$, $f(e^t, u) := \exp_x(ktI\mathcal{X}(x) + u)$. The differential of f at any point in $S^1 \times \{0\}$ is complex. Hence, if R is small enough, there exists an equivariant almost complex structure I_x on $\mathbb{C}^n$ satisfying $\|I_x - I_0\|_{L^\infty} < \epsilon$ and which induces a complex structure on $R \times \mathbb{C}^{n-1}$ with respect to which the restriction of f to a neighbourhood V of $R \times \mathbb{C}^{n-1}$ is complex. Also, since the restriction of f to $S_\rho \times \{0\}$ is an isometry, we can assume that the pullback metric f^*g extends to $\mathbb{C}^n$ satisfying $\|g - g_0\|_{L^\infty} < \epsilon$. Then we set $U = f(V)$ and $\xi := (f|_V)^{-1}$.

It is easy to see, using the compactness of X and the rigidity of representations of compact groups, that the set of weights of the representations which we construct as x moves along X forms a finite set. □

10.2. **An inequality for the energy on cylinders.** Let $C := [-2, 3] \times S^1$ with the standard product metric. Whenever we write any norm (L^∞, L^2, etc.) of either the connection or the section of a pair defined over C, unless we specify some other domain, we mean the norm over C. Define also the following subsets of C:

$$Z := [-1, 2] \times S^1, \quad Z_I := [-1, 2] \times S^1, \quad Z_{II} := [0, 1] \times S^1, \quad Z_{III} := [1, 2] \times S^1. \tag{10.35}$$

Theorem 10.5. *For any noncritical residue $\lambda \in \mathbf{i}\mathbb{R} \setminus \Lambda_{\mathrm{cr}}$ there exist real numbers $\epsilon = \epsilon(\lambda, X, I) > 0$ $\gamma = \gamma(\lambda, X, I) \in (0, 1/2)$, and $K = K(\lambda, X, I) > 0$, depending continuously on λ, such that if $(\alpha, \phi) : C \to X$ is an I-holomorphic pair satisfying the conditions $\|\alpha - \lambda d\theta\|_{L^\infty(C)} \le \epsilon$ and $\|d_\alpha\phi\|_{L^2(C)} \le \epsilon$, then*

(1) *the following inequality holds:*

$$\|d_\alpha\phi\|^2_{L^2(Z_{II})} \leq \gamma\left(\|d_\alpha\phi\|^2_{L^2(Z_I)} + \|d_\alpha\phi\|^2_{L^2(Z_{III})}\right);$$

(2) *if I, α and ϕ are of class C^1, then* $\sup_Z |d_\alpha\phi| \leq K\|d_\alpha\phi\|_{L^2(C)}$.

To prove Theorem 10.5 we need an analogous result for the case $X = \mathbb{C}^n$ which we now state. Fix a nontrivial diagonal action of S^1 on $\mathbb{C}^n$, and denote its weights by $w = (w_1, \ldots, w_n) \in \mathbb{Z}^n$. Consider on $\mathbb{C}^n$ the standard Riemannian metric g_0, and denote by I_0 the standard complex structure. Finally, define, for every $\eta \in \mathbb{R}$,

$$\gamma(\eta) := \frac{1}{e^{\eta/2} + e^{-\eta/2}}. \tag{10.36}$$

Lemma 10.6. *There exist real numbers $\epsilon = \epsilon(w) > 0$ and $K = K(w, \lambda)$ with the following property. Suppose that I is a smooth equivariant almost complex structure in $\mathbb{C}^n$ such that $\|I - I_0\|_{L^\infty} < \epsilon$, and that $(\alpha, \phi) : C \to \mathbb{C}^n$ is a I-holomorphic pair satisfying $\|\alpha - \lambda d\theta\|_{L^\infty(C)} \leq \epsilon$ and $\|d_\alpha\phi\|_{L^2(C)} \leq \epsilon$. Then*

(1) *the following inequality holds:*

$$\|d_\alpha\phi\|^2_{L^2(Z_{II})} \leq \gamma(2l_{\min})\left(\|d_\alpha\phi\|^2_{L^2(Z_I)} + \|d_\alpha\phi\|^2_{L^2(Z_{III})}\right), \tag{10.37}$$

where $l_{\min} := \min\{|-k + \mathbf{i}w_j\lambda| \mid k \in \mathbb{Z}, \quad 1 \leq j \leq n\}$;

(2) *we can bound:*

$$\sup_Z |d_\alpha\phi| \leq K\|d_\alpha\phi\|_{L^2(C)}. \tag{10.38}$$

Lemma 10.6 will be proved in Section 10.8 below. We now prove Theorem 10.5. Take ϵ smaller than the infimum of $\epsilon(w)$ as w moves along $W(X)$. Take, for this choice of ϵ and for each $x \in X$, an equivariant chart $\xi_x : U_x \to \mathbb{C}^n$ as in Lemma 10.4. Then take a finite subcovering of $\{U_x\}_{x\in X}$ and denote it by $\{U_1, \ldots, U_r\}$. Denote also by ξ_j and w_j the corresponding embeddings in $\mathbb{C}^n$ and weights. Since the sets U_j are S^1-invariant and open, it follows that there is a constant $\epsilon' > 0$ such that for any $K \subset X$ satisfying $\operatorname{diam}_{S^1} K < \epsilon'$ there is at least one j such that $K \subset U_j$. Now, if we take $\delta > 0$ small enough, it follows from Lemma 10.14 that $\operatorname{diam}_{S^1} \phi(C) < \epsilon'$. Then the statement of Theorem 10.5 follows from Lemma 10.6 applied to $\xi_j \circ \phi : C \to \mathbb{C}^n$. Indeed, the condition that λ is not critical implies that, applying Lemma 10.6 to any chart, we have that $l_{\min} > 0$, so $\gamma(2l_{\min}) < 1/2$. Then γ can be defined as the maximum of the numbers $\gamma(2l_{\min})$ computed for each of the charts $U_1, \ldots, U_r$.

10.3. **From mean inequalities to exponential decay.** The following lemma will allow to obtain exponential decay from inequalities of the type given by (1) in Lemma 10.6. We will use it several times in this paper.

Lemma 10.7. *Let $\{x_0, \ldots, x_N\}$ be a sequence of nonnegative real numbers satisfying, for some $\gamma \in (0, 1/2)$ and each k between 1 and $N - 1$,*

$$x_k \leq \gamma(x_{k+1} + x_{k-1}). \tag{10.39}$$

Let $\xi := (1+\sqrt{1-4\gamma^2})/(2\gamma)$. Then $\xi > 1$ and for any k between 0 and N we have

$$x_k \leq x_0\xi^{-k} + x_N\xi^{-(N-k)}. \tag{10.40}$$

Proof. That $\xi > 1$ is an easy computation. As for (10.40), observe that $\xi^{-1} = \gamma(1+\xi^{-2})$. Consequently, if we define $y_k := x_k - (x_0\xi^{-k} + x_N\xi^{-(N-k)})$ then $y_k \leq \gamma(y_{k+1} + y_{k-1})$. Since $\gamma \in (0, 1/2)$, it follows that the sequence $\{y_k\}$ attains its maximum at y_0 or y_N. Both of these numbers are ≤ 0, so all y_j are nonpositive. This proves (10.40). □

Corollary 10.8. *Suppose that $\{x_0, x_1, x_2, \dots\}$ is an infinite sequence of nonnegative real numbers satisfying $x_k \leq \gamma(x_{k+1} + x_{k-1})$ for some $\gamma \in (0, 1/2)$ and any $k \geq 1$. Suppose that $x_j \to 0$ as j goes to infinity. Define $\xi > 1$ as in Lemma 10.7. Then for any k we have $x_k \leq x_0\xi^{-k}$.*

Proof. Apply Lemma 10.7 to the first N terms and make N go to infinity. □

Lemma 10.9. *Suppose that $\{x_{-N}, x_{-N+1}, \dots, x_N\}$ and $\{z_{-N}, x_{-N+1}, \dots, z_N\}$ are sequences of nonnegative real numbers. Assume that there are positive constants $\gamma \in (0, 1/2)$, χ, K and ϵ satisfying:*

- *for every j, $z_j \leq Ke^{-\chi(N-|j|)}$;*
- *for every $-N+2 \leq j \leq N-2$, if $z_{j-2} + \dots + z_{j+2} \leq \epsilon(x_{j-2} + \dots + x_{j+2})$ then $x_j \leq \gamma(x_{j-1} + x_{j+1})$.*

Let $\xi = \xi(\gamma)$ be as in Lemma 10.7, and let $\sigma := \min\{\chi, \ln\xi\}$. Then we have, for any $-N+1 \leq j \leq N-1$,

$$x_j \leq (\epsilon^{-1}10e^{2\chi}K + x_{-N+1} + x_{N-1})e^{-\sigma(N-|j|)}. \tag{10.41}$$

Proof. For any $-N+2 \leq j \leq N-2$ we say that x_j is big if $x_j \geq \epsilon^{-1}(z_{j-2} + \dots + z_{j+2})$, and otherwise we say that x_j is small. We have $z_{j-2} + \dots + z_{j+2} \leq 5e^{2\chi}Ke^{-\chi(N-|j|)}$. In particular, if x_j is small then $x_j \leq \epsilon^{-1}5e^{2\chi}Ke^{-\chi(N-|j|)}$. Take now any x_j. If x_j is small then the previous inequality implies (10.41). So we assume that x_j is big and we take the longest sequence of consecutive big elements $x_{p+1}, \dots, x_j, \dots, x_{q-1}$ containing x_j. Suppose that x_p and x_q are small (this need not be the case, since we could have $p = -N+1$ or $q = N-1$). By Lemma 10.7 we have $x_j \leq x_pe^{-\sigma(j-p)} + x_qe^{-\sigma(q-j)}$. Now, since both x_p and x_q are small we can bound (using $\sigma \leq \chi$)

$$x_j \leq \epsilon^{-1}5e^{2\chi}K(e^{-\sigma(N-|p|+j-p)} + e^{-\sigma(N-|q|+q-j)}) \leq \epsilon^{-1}10e^{2\chi}K(e^{-\sigma(-N+|j|)}),$$

because both $N-|p|+j-p$ and $N-|q|+q-j$ are greater than $N-|j|$. When $p = -N+1$ or $q = N-1$ we proceed similarly. □

10.4. **Proof of Theorem 10.1.** By Lemma 3.1 (see Remark 3.4) we can trivialize P in such a way that $d_A = d + \lambda d\theta + \alpha'$, where $\lambda = \operatorname{Res}(A, 0, \tau)$ and α' is a continuous 1-form on $\mathbb{D}$. Let $\alpha = \lambda d\theta + \alpha'$. We distinguish two situations, according to whether the residue is critical or not.

Suppose first that the residue is critical. In particular, $\mathbf{i}\lambda$ is rational, so we can write $\lambda = \mathbf{i}p/q$, where p, q are relatively prime integers and $q \neq 1$. Let $\pi : \mathbb{D}^* \to \mathbb{D}^*$ be the map

given by $\pi(z) := z^q$. Then $d_{\pi^*A} = d + q\lambda d\theta + \pi^*\alpha' = d + pd\theta + \pi^*\alpha'$. Let $g : \mathbb{D}^* \to S^1$ be defined by $g(r,\theta) := e^{2\pi \mathbf{i} p\theta}$. Then

$$d_B := d_{g^*\pi^*A} = d + \pi^*\alpha'$$

is a continuous connection on $\mathbb{D}$. Let $I(B)$ be the continuous almost complex structure on $\mathbb{D} \times X$ induced by B (see Section 2). By (2.8), the map $\Phi = (\iota, g \cdot (\phi \circ \pi)) : \mathbb{D}^* \to \mathbb{D} \times X$ is $I(B)$-holomorphic (here $\iota : \mathbb{D}^* \to \mathbb{D}$ denotes the inclusion), and by (2.9) $d\Phi$ has bounded L^2 norm. Hence we can apply the theorem on removal of singularities for holomorphic curves (as proved for continuous almost complex structures in Corollary 3.6 of [IS]), deduce that Φ extends to a $I(B)$-holomorphic map $\Phi : \mathbb{D} \to \mathbb{D} \times X$, and consequently obtain an extension $g \cdot (\phi \circ \pi) : \mathbb{D} \to X$. Furthermore, this extension is of type C^1 (because the complex structure is continuous). Let now $\zeta := e^{2\pi \mathbf{i}/q}$. Then $(g\cdot(\phi\circ\pi))(\zeta z) = (g\cdot(\phi\circ\pi))(z)$, so that $d(g\cdot(\phi\circ\pi))(0) = 0$. Furthermore, since $d\pi(0) = 0$, $\pi^*\alpha'$ vanishes at 0. It follows from this that $d_{g^*\pi^*A}(g\cdot(\phi\circ\pi)) = (d+\pi^*\alpha')(g\cdot(\phi\circ\pi)) = 0$, from which we deduce that $|d_{g^*\pi^*A}(g \cdot (\phi \circ \pi))(r,\theta)| = |d_{\pi^*A}(\phi \circ \pi)(r,\theta)| < Kr$ for some constant K. This then implies that $|d_A\phi(r,\theta)| < K' r^{1/q}$ for some other constant K', so the claim is proved.

Now consider the case of noncritical residue $\lambda \notin \Lambda_{\mathrm{cr}}$. Consider the cylinder $\mathbb{R}^+ \times S^1$ as a conformal model of $\mathbb{D}^*$, with coordinates $t \in \mathbb{R}^+$ and $\theta \in S^1$ and the standard flat metric $dt^2 + d\theta^2$. We have:

$$|\alpha'(t,\theta)| < Ke^{-t} \qquad \text{and} \qquad \|d_\alpha\phi\|_{L^2} < \infty. \tag{10.42}$$

We look at ϕ as a map from $\mathbb{D}^*$ to X, so that (α,ϕ) is a pair. Define, for any $n \in \mathbb{N}$, $Z_n = [n, n+1] \times S^1$ and let $f_n := \|d_\alpha\phi\|^2_{L^2(Z_n)}$. We claim that, for big enough n, we have

$$f_{n+1} \le \gamma(\lambda, X)(f_n + f_{n+2}), \tag{10.43}$$

where $\gamma(\lambda, X)$ is as in Theorem 10.5. Indeed, thanks to (10.42), if n is big enough then both the L^∞ norm of the restriction of α' to $[n, n+3] \times S^1$ and the L^2 norm of $d_\alpha\phi$ restricted to $[n, n+3] \times S^1$ are less than the value of ϵ given by Theorem 10.5. Hence, we can apply (1) in Theorem 10.5 (identifying $Z_I = Z_n$, $Z_{II} = Z_{n+1}$ and $Z_{III} = Z_{n+2}$) and deduce (10.43). Combining Corollary 10.8 with (10.43) we deduce that

$$\|d_\alpha\phi\|^2_{L^2(Z_n)} = f_n \le K\xi^{-n},$$

where $\xi > 1$. Now, using (2) in Theorem 10.5 we deduce the pointwise bound

$$|d_\alpha\phi(t,\theta)| \le K'\xi^{-t}$$

for some other constant K'. This finishes the proof of the case of noncritical residue.

10.5. **Proof of Theorem 10.3.** The proof is exactly like that of the case of noncritical residue of Theorem 10.1, in Section 10.4 above. Namely, we define f_n to be $\|d_\alpha\phi\|_{L^2([n,n+1]\times S^1)}$ and we use Theorem 10.5 (provided ϵ has been chosen small enough) to deduce that $f_{n+1} \le \gamma(f_n + f_{n+2})$ for some $\gamma \in (0, 1/2)$ depending on λ. Then Lemma 10.7 gives $f_n \le Ke^{-\sigma(N-|n|)}(f_{-N} + f_{N-1})$ for some K and σ. To deduce the pointwise bound (10.33) we use (2) in Theorem 10.5. Finally, (10.34) follows easily from (10.33) combined with Lemma 2.1.

10.6. **Some convexity properties of holomorphic maps.** In this subsection we state a result which will be crucial for proving most of the convexity results in this paper. Part of it can be seen as a particular case of Lemma 10.6 (concretely, the case $I = I_0$ and $\alpha' = 0$).

For any $\delta \in (0, 1/2)$ we will denote

$$Z^\delta := [-1+\delta, 2-\delta], \qquad Z_I^\delta := Z_I \cap Z^\delta, \qquad Z_{III}^\delta := Z_{III} \cap Z^\delta. \tag{10.44}$$

Lemma 10.10. *Take on $\mathbb{C}^n$ the standard almost complex structure I_0 and metric g_0, and suppose that S^1 acts diagonally on $\mathbb{C}^n$ with weights $w_1, \dots, w_n$. Let $\lambda \in \mathbf{i}\mathbb{R}$ and let $\phi : C \to \mathbb{C}^n$ be a smooth map satisfying $\overline{\partial}_{\lambda d\theta}\phi = 0$. Let*

$$l_{\min} := \min\{|-k + \mathbf{i}w_j\lambda| \mid k \in \mathbb{Z},\ 1 \le j \le n\}.$$

Let γ be the function defined in (10.36).

1. *If $l_{\min} \neq 0$ then we have, for some $\delta > 0$ depending on $l_{\min}$ but independent of ϕ:*
$$\|d_{\lambda d\theta}\phi\|^2_{L^2(Z_{II})} < \gamma(2l_{\min})(\|d_{\lambda d\theta}\phi\|^2_{L^2(Z_I^\delta)} + \|d_{\lambda d\theta}\phi\|^2_{L^2(Z_{III}^\delta)}).$$
2. *Suppose that $\lambda = 0$ (so that ϕ is holomorphic) and define*
$$\phi_{\mathrm{av}}(t,\theta) := \phi(t,\theta) - \frac{1}{2\pi}\int \phi(t,\nu)d\nu.$$
(Here the subscript av *stands for average.) Then, if ϕ is not constant, the following holds:*
$$\|\phi_{\mathrm{av}}\|^2_{L^2(Z_{II})} \le \frac{1}{e^2+e^{-2}}(\|\phi_{\mathrm{av}}\|^2_{L^2(Z_I)} + \|\phi_{\mathrm{av}}\|^2_{L^2(Z_{III})}).$$

Proof. The condition $\overline{\partial}_{\lambda d\theta}\phi = 0$ is equivalent to

$$\frac{\partial\phi}{\partial t} = I_0\left(\frac{\partial\phi}{\partial\theta} - \mathbf{i}\lambda\mathcal{X}(\phi)\right), \tag{10.45}$$

where, denoting by $(\phi_1, \dots, \phi_n)$ the coordinates of ϕ, we have $\mathcal{X}(\phi) = (\mathbf{i}w_1\phi_1, \dots, \mathbf{i}w_n\phi_n)$. Writing for each j the Fourier expansion $\phi_j(t,\theta) = \sum_{k\in\mathbb{Z}} a_{k,j}(t)e^{\mathbf{i}k\theta}$ formula (10.45) is equivalent to the equation $a'_{k,j} = (-k + \mathbf{i}\lambda w_j)a_{k,j}$ for each Fourier coefficient. Hence $a_{k,j}(t) = a_{k,j}(0)e^{(-k+\mathbf{i}\lambda w_j)t}$. Also, since $\overline{\partial}_{\lambda d\theta}\phi = 0$ we have $|d_{\lambda d\theta}\phi|^2 = 2|\partial\phi/\partial t|^2$. Putting everything together we write

$$\int_{[t_0,t_1]\times S^1} |d_{\lambda d\theta}\phi|^2 = \sum_{j=1}^n \sum_{\substack{k\in\mathbb{Z}\\ \mathbf{i}k+\lambda\neq 0}} \int_{t_0}^{t_1} 2|a_{k,j}(0)|^2|\mathbf{i}k+\lambda|^2 e^{2(-k+\mathbf{i}\lambda w_j)t}dt. \tag{10.46}$$

The proof now follows from Lemma 10.11 below. This proves (1). The formula in (2) is proved similarly, noting that the integral which computes $\|\phi_{\mathrm{av}}\|^2$ is of the form $\int \sum_{j\in\mathbb{Z}\setminus\{0\}} |a_{k,j}|^2 e^{2jt}dt$. □

Lemma 10.11. *For any $\eta_0 > 0$ there is some $\delta > 0$ such that for every $\eta \in \mathbb{R}$ satisfying $|\eta| \geq \eta_0$ we have*

$$\int_0^1 e^{\eta x} dx < \gamma(\eta_0) \left(\int_{-1+\delta}^0 e^{\eta x} dx + \int_1^{2-\delta} e^{\eta x} dx \right). \tag{10.47}$$

Proof. By symmetry it suffices to consider the case $\eta > 0$. A simple computation shows that the inequality

$$\int_0^1 e^{\eta x} dx < \beta \left(\int_{-1}^0 e^{\eta x} dx + \int_1^2 e^{\eta x} dx \right) \tag{10.48}$$

is equivalent to $\beta > 1/(e^{\gamma} + e^{-\gamma})$. Since the function $f(x) := e^x + e^{-x}$ is increasing for positive x, if we set $\beta := \gamma(\eta_0/2)$ then (10.48) holds for every $\eta \geq \eta_0$. On the other hand, if we set $\delta := 1/2$ then (10.47) holds for every η bigger than some $\eta' > \eta_0 > 0$. Indeed, the left hand side in (10.47) grows as a function of η as e^{η}, whereas the right hand side grows as $e^{\eta(2-\delta)} = e^{3\eta/2}$. On the other hand, since the inequality (10.48) is strict for $\beta = \gamma(\eta_0)$ and every $\eta \geq \eta_0$, we can pick some very small $0 < \delta < 1/2$ such that (10.47) holds for every $\eta \in [\eta_0, \eta]$. Then (10.47) will also hold for every $\eta \geq \eta_0$ (of course, if for a given η (10.47) is true for some $\delta = \delta_0$, then it is also true for any $0 \leq \delta \leq \delta_0$). □

10.7. **Compactness of the set of pairs of small energy.** A sequence of pairs $\{(\alpha_n, \phi_n) : C \to X\}$ will be said to **converge** to a pair $(\alpha, \phi) : C \to X$ if $\alpha_n \to \alpha_0$ in the L^∞ norm and $\phi_n \to \phi_0$ in $L^p_{1,\mathrm{loc}}$ for any $p < \infty$. This implies that for any compact subset $K \subset C$ the energies $\|d_{\alpha_n}\phi_n\|_{L^2(K)}$ converge to $\|d_{\alpha_0}\phi_0\|_{L^2(K)}$, and that the maps ϕ_n converge to ϕ_0 in the continuous topology.

The following lemma is, except for the last statement, a combination of Lemma 3.1 and Corollary 3.3 in [IS] (see Definition 3.1 in [IS] for the notion of uniformly continuous almost complex structure on a manifold).

Lemma 10.12. *Let Y be a (nonnecessarily compact) manifold. Let h be some metric on Y, and let I_0 be a continuous almost complex structure on Y. Let $Z \subset Y$ be a closed h-complete subset, such that I_0 is uniformly continuous on Z w.r.t. h. There exists a real number $\epsilon = \epsilon(I_0, Z, h)$ with the following property. Let $\{I_n\}$ be a sequence of continuous almost complex structures on Y such that $I_n \to I_0$ in C^0-topology of Y. Let $u_n \in C^0 \cap L^2_{1,\mathrm{loc}}(\mathbb{D}, Y)$ be a sequence of I_n-holomorphic maps such that $u_n(\mathbb{D}) \subset Z$, $\|du_n\|_{L^2(\mathbb{D})} \leq \epsilon$ and $u_n(0)$ is bounded in X. Then there exists a subsequence $\{u_{n_k}\}$ which $L^p_{1,\mathrm{loc}}$ converges to a I_0-holomorphic map u_∞ for all $p < \infty$. In particular, for any $K \Subset \mathbb{D}$ the norms $\|du_n\|_{L^2(K)}$ tend to $\|du_\infty\|_{L^2(K)}$. Furthermore, if for some $p > 2$ and $k > 0$ the L^p_k norms of I_n are uniformly bounded and $\|I_n - I_0\|_{L^p_k} \to 0$, then u_∞ is of class L^p_k and the subsequence can be chosen to converge to u_∞ in L^p_k.*

Proof. We prove the last statement: by Proposition B.4.7 in [McDS], each ϕ_n is of class L^p_{k+1}, and since $L^p_{k+1} \to L^p_k$ is compact, passing to a subsequence there is a limit $\phi_n \to \phi$ in L^p_k. □

Using the previous lemma we prove the following result on convergence of pairs with small energy.

Lemma 10.13. *Let $\{I_n\}$ be a sequence of continuous almost complex structures on X such that $I_n \to I$ in the L^∞ norm, where I is also a continuous almost complex structure. Let C be a (nonnecessarily compact) complex curve with a Riemannian metric. Let $(\alpha_n, \phi_n) : C \to X$ be a sequence of I_n-holomorphic pairs. Assume that $\|d\alpha_n\|_{L^\infty}$ is uniformly bounded, and that $\|d_{\alpha_n}\phi_n\|_{L^2(C)} \leq \epsilon/2$ (where ϵ is as in Lemma 10.12). Then, for any compact subset $M \subset C$, there exists a subsequence of pairs $\{(\alpha_{n_k}, \phi_{n_k})\}$ whose restriction to M converges, up to regauging, to a I-holomorphic pair $(\alpha, \phi) : M \to X$. If for some $p > 2$ and $k > 0$ the L^p_k norms of α_n, I_n, α and I are uniformly bounded, then ϕ is of class L^p_k and the subsequence can be taken so that $\phi_n \to \phi$ in L^p_k.*

Proof. Since $d\alpha_n$ has uniformly bounded L^∞ norm we can assume, up to regauging, that the sequence α_n is uniformly continuous and L^∞ bounded. By Ascoli–Arzela it follows that there is a subsequence which converges in L^∞ norm to a continuous 1-form $\alpha \in \Omega^1(D, \mathbf{i}\mathbb{R})$. Take the corresponding sequence of pairs and denote it again by $\{(\alpha_n, \phi_n)\}$, so that α_n converges to α in the L^∞ norm as $n \to \infty$. By a standard patching argument in gauge theory (see §4.4.2 in [DK]) it suffices to prove the result of the lemma for a finite collection of disks covering C. Let h_0 be the product metric on $Y = C \times X$, let $\Phi_n : C \to Y$ be the section corresponding to ϕ_n, and let $h_n := g(\alpha_n)$ and $h := g(\alpha)$. Then we have (see (2.9)) $|d\Phi_n|^2_{h_n} = 1 + |d_{\alpha_n}\phi_n|^2$. Since $\alpha_n \to \alpha$ in L^∞ we deduce that for a big enough n we can bound $h_n < \sqrt{2}h$. It follows that, if $D \Subset \mathbb{D}$ is a disk of area $\epsilon^2/(\sqrt{2}K)$, where K is a suitable constant depending on the metrics h_0 and h, then $\|d\Phi_n\|_{h,L^2(D)} \leq \epsilon$. Hence, if $f : \mathbb{D} \to D$ is an affine biholomorphism, then $\|d(\Phi_n \circ f)\|_{h,L^2(\mathbb{D})} \leq \epsilon$. Finally, it follows from $\alpha_n \to \alpha$ that $I_n(\alpha_n) \to I(\alpha)$ in L^∞ norm. Taking into account (2.8), we can apply Lemma 10.12 (it is clear that $Z := D \times X$ satisfies the hypothesis of Lemma 10.12) to the sequence of maps $\{\Phi_n \circ f\}$, and deduce that a subsequence converges to a map defined on D. The last statement follows from the last statement in Lemma 10.12 (since $p > 2$ and $k > 1$, if α_n has bounded L^p_k norm then $I(\alpha_n)$ is also bounded in L^p_k). □

In the following two corollaries we use the same notation as in Section 10.2, so C denotes the cylinder $[-1, 2] \times S^1$ and $Z = [0, 1] \times S^1$.

Corollary 10.14. *Let I be an almost complex structure on X. For any $\epsilon > 0$ there exists some $\delta > 0$ such that if $(\alpha, \phi) : C \to X$ is a I-holomorphic pair satisfying $\|d\alpha\|_{L^\infty(C)} < \delta$ and $\|d\phi\|_{L^2(C)} < \delta$ then*

$$\operatorname{diam}_{S^1} \phi(Z) < \epsilon.$$

Proof. We prove the corollary by contradiction. Take any $\epsilon > 0$ and assume that there is a real number $\delta > 0$ and a sequence of I-holomorphic pairs $(\alpha_n, \phi_n) : C \to X$ such that $\|d\alpha_n\|_{L^\infty(C)} \to 0$, $\|d_{\alpha_n}\phi_n\|_{L^2(C)} \to 0$ and $\operatorname{diam}_{S^1} \phi_n(Z) \geq \delta$. By Lemma 10.13 we can assume, up to restricting to a subsequence and regauging, that there is a I-holomorphic pair $(\alpha, \phi) : Z \to X$ such that the restriction of $\{(\alpha_n, \phi_n)\}$ to Z converges to (α, ϕ).

This implies in particular that $\|d_\alpha\phi\|_{L^2(Z)} = 0$ and that $\operatorname{diam}_{S^1}\phi(Z) \geq \delta$, which is impossible. □

Corollary 10.15. *For any $p > 2$, K and $\epsilon > 0$ there is some $\delta > 0$ with the following property. Suppose that $(\alpha, \phi) : Z \to X$ is a I-holomorphic pair, that $\|d\alpha\|_{L^p_1} \leq K$, and that $\sup_{\{1/2\}\times S^1} |d_\alpha\phi| \geq \epsilon$. Then $\|d_\alpha\phi\|_{L^2(Z)} \geq \delta$.*

Proof. Fix p, K and ϵ, and suppose there is no $\delta > 0$ satisfying the hypothesis of the Corollary. Then there is a sequence of pairs $(\alpha_u, \phi_u) : Z \to X$ such that $\|d\alpha_u\|_{L^p_1} \leq K$, $\sup_{\{1/2\}\times S^1} |d_{\alpha_u}\phi_u| \geq \epsilon$ and $\|d_{\alpha_u}\phi_u\|_{L^2(Z)} \to 0$. Using the last statement in Lemma 10.13, we deduce that there is a subsequence (which we denote again by (α_u, ϕ_u)) which converges to (α, ϕ) in L^p_2. In particular, $d_\alpha\phi = 0$, but also $d_{\alpha_u}\phi_u$ converges pointwise to $d_\alpha\phi$, which is a contradiction. □

10.8. **Proof of Lemma 10.6.** The strategy will be to reduce the lemma to the case of the standard complex structure I_0 on $\mathbb{C}^n$ and $\alpha = \lambda d\theta$. We will do this by means of a compactness argument. Before stating the argument, we need some preliminaries. Permuting the coordinates if necessary, we can assume that for some p we have

$$\text{for any } 1 \leq j \leq p,\ w_j \neq 0 \qquad \text{and} \qquad w_{p+1} = \cdots = w_n = 0. \tag{10.49}$$

Let $q = n - p$, so that we have a splitting $\mathbb{C}^n = \mathbb{C}^p \times \mathbb{C}^q$. Let π_p, π_q be the projections from $\mathbb{C}^n$ to $\mathbb{C}^p$ and $\mathbb{C}^q$. The following lemma will be proved in Section 10.9.

Lemma 10.16. *There exists some constant $K > 0$ with the following property. Let $(\alpha, \phi) : Z \to \mathbb{C}^n$ be a I-holomorphic pair such that $\|\alpha - \lambda d\theta\|_{L^\infty} < |\lambda|/2$. For any $0 < \delta < 1/2$ there exists some $y \in \mathbb{C}^q$ such that*

$$\sup_{z\in Z^\delta} |\phi(z) - (0, y)| \leq l_{\min}^{-1}\delta^{-1/2}K(\|d_\alpha\phi\|_{L^2} + 1).$$

We first prove (10.37). Suppose that no matter how small ϵ is there are pairs for which (10.37) does not hold. Then one can chose a sequence of equivariant almost complex structures I_u and I_u-holomorphic pairs $(\alpha_u, \phi_u) : C \to \mathbb{C}^n$ satisfying:

$$\|I_u - I_0\|_{L^\infty} \to 0, \qquad \|\alpha_u - \lambda d\theta\|_{L^\infty} \to 0, \qquad \|d_{\alpha_u}\phi_u\|_{L^2(C)} \to 0,$$

and for which the inequality (10.37) holds in the opposite direction:

$$\|d_\alpha\phi\|^2_{L^2(Z_{II})} > \gamma(2l_{\min})\left(\|d_\alpha\phi\|^2_{L^2(Z_I)} + \|d_\alpha\phi\|^2_{L^2(Z_{III})}\right). \tag{10.50}$$

In particular, the energy E_u of (α_u, ϕ_u) is nonzero. Let ϵ be less that the epsilons appearing in Lemmae 10.16 and 10.13, and define for every n a new pair (α'_u, ϕ'_u) by setting $\alpha'_u := \alpha_u$ and $\phi'_u(z) := \phi_u(z)\epsilon/2E_u$. Let also I'_u be the pullback of I_u by the homotecy of ratio $2E_u/\epsilon$. Then (α'_u, ϕ'_u) is a I'_u-holomorphic pair and its energy is $\epsilon'/2$. Furthermore, since I_0 is invariant under homotecies, it follows that $\|I'_u - I_0\|_{L^\infty} \to 0$.

Let $\delta > 0$ be the number given by (1) in Lemma 10.10 for our value of $l_{\min}$. Let $y_u \in \mathbb{C}^q$ a point, as given by Lemma 10.16, such that $\phi'_n(Z^\delta)$ is contained in the ball centered at $(0, y)$ and of radius $Kl_{\min}^{-1}\delta^{-1}$. Finally, let $\phi''_u := \phi'_u - (0, y_u)$ and let I''_u be the

pullback of I'_u by the translation in the direction $(0, y_u)$. It follows that (α'_u, ϕ''_u) is I''_u-holomorphic (note that I''_u is S^1-equivariant) and $\phi''_u(Z_\delta)$ is contained in the ball centered at 0 of radius $(\epsilon+1)Kl_{\min}^{-1}\delta^{-1/2}$. At this point we use Lemma 10.13 to deduce that, up to regauging, there is a subsequence of $\{(\alpha'_u, \phi''_u)\}$ which converges to a I_0-holomorphic pair $(\lambda d\theta, \phi) : Z^\delta \to \mathbb{C}^n$ which has energy $\epsilon/2$ and such that

$$\|d_{\lambda d\theta}\phi\|_{L^2(Z_{II})} \geq \gamma(2l_{\min})(\|d_{\lambda d\theta}\phi\|_{L^2(Z_I^\delta)} + \|d_{\lambda d\theta}\phi\|_{L^2(Z_{III}^\delta)}). \tag{10.51}$$

Now, comparing this with (1) in Lemma 10.10 we obtain a contradiction. This finishes the proof of (10.37).

The inequality (10.38) can be reduced, following the same strategy, to the case of the standard complex structure I_0 on $\mathbb{C}^n$ and the connection $\alpha = \lambda d\theta$. Suppose for simplicity that $n = 1$ and that $w_1 = 1$ (the general case offers no extra difficulty). Let $\phi : C \to \mathbb{C}$ be a map which satisfies

$$\frac{\partial\phi}{\partial t} = \mathbf{i}\left(\frac{\partial\phi}{\partial\theta} + \lambda\phi\right). \tag{10.52}$$

We want to prove that

$$\sup_Z \left|\frac{\partial\phi}{\partial t}\right| \leq K \left\|\frac{\partial\phi}{\partial t}\right\|_{L^2(C)}$$

for some constant K. Let $\psi := \partial\phi/\partial t$. Then ψ satisfies the same equation (10.52) as ϕ. It follows from that that the function $\zeta = e^{-\mathbf{i}\lambda t}\psi$ is holomorphic on C. Define the constants $K_0 := \sup_{t\in[0,1]} e^{\mathbf{i}\lambda t}$ and $K_1 := \sup_{t\in[0,1]} e^{-\mathbf{i}\lambda t}$. Take any $z \in Z$ and denote by $\mathbb{D} \subset C$ the disk of radius 1 centered at z. Since ζ is holomorphic we have $|\zeta(z)| \leq \|\zeta\|_{L^2(\mathbb{D})}$. Now we bound:

$$|\psi(z)| \leq K_0|\zeta(z)| \leq K_0\|\zeta\|_{L^2(\mathbb{D})} \leq K_0K_1\|\psi\|_{L^2(\mathbb{D})} \leq K_0K_1\|\psi\|_{L^2(C)},$$

which is what we wanted to prove.

10.9. **Proof of Lemma 10.16.** First of all we state and prove two lemmae which will be necessary for proving Lemma 10.16. For the first one we introduce the following notation. Let M be the manifold $\mathbb{R} \times S^1 \times \mathbb{C}^n$. Take a 1-form α on $\mathbb{R} \times S^1$ with values in $\mathbf{i}\mathbb{R}$, and let $J := I(\alpha)$ be the complex structure on M induced by α and the action of S^1 on $\mathbb{C}^n$ with weights $w_1, \dots, w_n$ as in (10.49). Consider also the metric $h := g(\alpha)$ on M induced by α and the standard metric g_0 on $\mathbb{C}^n$.

Lemma 10.17. *For any $K' > 0$ there exists some $K'' > 0$ with the following meaning. Let $\alpha \in \Omega^1(\mathbb{R} \times S^1, \mathbf{i}\mathbb{R})$ be a connection 1-form. Suppose that $\|\alpha\|_{L^\infty} \leq K'$ and that $\Phi = (i, \phi) : Z \to M$ is a J-holomorphic map, where $i : Z \to \mathbb{R} \times S^1$ is the inclusion; suppose also that for some $z \in Z$ and $R > 0$ we have $d(\phi(z), \phi(\partial Z)) > R$. Then*

$$\|d\Phi\|_{L^2}^2 = \operatorname{area}(Z) + \|d_\alpha\phi\|_{L^2}^2 > K''R^2.$$

Proof. This is similar to a standard result in the theory of pseudo-holomorphic curves. First, since Φ is J-holomorphic, $\|d\Phi\|^2$ is equal to the area of $\Phi(Z)$, so we have to prove

that $\mathrm{area}(\Phi(Z)) > K'' R^2$. Define, for any $R \geq r > 0$, the following sets:

$$M_r = \{(t, \theta, x) \in M \mid |x - \phi(z)| \leq r\}, \qquad B_r = \partial M_r, \qquad Z_r = \Phi^{-1}(M_r).$$

Define also $f(r) := \mathrm{area}(\Phi(Z_r))$. A real number r will be called regular if it is a regular value of the function $Z \ni y \mapsto |\phi(y) - \phi(z)|$. For any such r, $\Phi(Z_r) \subset M$ is a smooth subsurface with boundary. Also, the set of nonregular values has measure zero.

We first prove an upper bound for $f(r)$. There exists some constant $K_1 > 0$ such that for any regular r there is a smooth subsurface $\Sigma \subset M$ such that $\partial\Sigma = \partial\Phi(Z_r)$ and such that $\mathrm{area}_h(\Sigma) \leq K_1 \,\mathrm{length}(\partial\Phi(Z_r))^2$. Let $\omega(\alpha)$ be the minimal coupling symplectic on M (see Section 2). Since $H^2(M, \mathbb{R}) = 0$ and since $\omega(\alpha)$ has bounded L^∞ norm (because α is bounded), it follows that, for some constant K_2 and any regular r,

$$\int_{\Phi(Z_r)} \omega(\alpha) = -\int_\Sigma \omega(\alpha) \leq K_2 \,\mathrm{area}(\Sigma) \leq K_2 K_1 \,\mathrm{length}(\partial\Phi(Z_r))^2.$$

On the other hand, since Φ is $I(\alpha)$-holomorphic we can identify the first integral with the area of $\Phi(Z_r)$. Hence we have proved

$$f(r) \leq K_2 K_1 \,\mathrm{length}(\partial\Phi(Z_r))^2. \tag{10.53}$$

To finish the proof of the lemma, observe that there exists some constant K_3 such that for any regular r we have $f'(r) \geq K_3 \,\mathrm{length}(\partial\Phi(Z_r))$. Combining this with (10.53) we deduce $f'(r)^2 \geq K_3 f(r)/(K_2 K_1)$. Integrating over r we deduce the result. □

Let $\mathcal{X} : \mathbb{C}^n \to \mathbb{C}^n$ be the vector field generated by the infinitesimal action of S^1 on $\mathbb{C}^n$.

Lemma 10.18. *There exists some constant K_1, independent of λ, such that if $f : S^1 \to \mathbb{C}^p \subset \mathbb{C}^n$ is any smooth map, and $\beta : S^1 \to \mathbf{i}\mathbb{R}$ satisfies $|\beta - \lambda| < \lambda/2$ then*

$$\sup |f| \leq l_{\min}^{-1} K_1 \|f' + \beta\mathcal{X}(f)\|_{L^2(S^1)}.$$

Proof. Let $\beta_0 := \frac{1}{2\pi}\int \beta(\theta) d\theta$. Define $g := e^s f$, where $s : S^1 \to \mathbf{i}\mathbb{R}$ satisfies $ds = \beta - \beta_0$. We have $|f| = |g|$, and one checks easily that $|f' + \beta\mathcal{X}(f)| = |g' + \beta_0\mathcal{X}(g)|$ pointwise. Hence it suffices to prove that for some universal constant K_1 (depending on λ) and any constant β_0 satisfying $|\beta_0 - \lambda| < \lambda$ one has the inequality

$$\sup |g| \leq l_{\min}^{-1} K_1 \|g' + \beta_0\mathcal{X}(g)\|_{L^2(S^1)}.$$

Now, this can be proved by considering the Fourier series of g. □

After these preliminaries, we now prove Lemma 10.16. Pick K_1 satisfying the requirement of the previous theorem and in such a way that for any $h : S^1 \to \mathbb{C}^q \subset \mathbb{C}^n$ and any $\theta, \eta \in S^1$ we have

$$|h(\theta) - h(\eta)| \leq K_1 \|h'\|_{L^2(S^1)} \tag{10.54}$$

(this is possible in view of Cauchy–Schwartz). Let K'' be the constant given by Lemma 10.17 for the value $K' = 3|\lambda|/2$. Take $K > 0$ big enough so that for any positive $E > 0$

we have

$$l_{\min}^{-1}\delta^{-1/2}K(E+1) > 6\max\left\{l_{\min}^{-1}\delta^{-1/2}K_1E,\sqrt{(E^2+2\pi)/K''}\right\}. \tag{10.55}$$

There exists some $a\in[0,\delta]$ such that

$$\|d_\alpha\phi\|^2_{L^2(\{a\}\times S^1)} \le \delta^{-1}\|d_\alpha\phi\|^2.$$

Denote the components of $\phi(a,\theta)$ by $(f(\theta),h(\theta))$. Let $\beta: S^1\to\mathbf{i}\mathbb{R}$ be the map such that for any θ we have $\alpha(a,\theta)=\beta(\theta)d\theta+\gamma(\theta)dt$. Applying Lemma 10.18 to the component f and inequality (10.54) to the component h (together with $l_{\min}<1$) we deduce that

$$\sup_{\theta\in S^1}|\phi(a,\theta)-(0,y_a)| \le \rho := l_{\min}^{-1}\delta^{-1/2}K_1\|d_\alpha\phi\|_{L^2},$$

where $y_a=h(1)$ (indeed, $f'+\beta\mathcal{X}(f)$ can be identified with $d_\alpha\pi_p(\phi)$ to the circle $\{a\}\times S^1$, and we have $\|d_\alpha\pi_p(\phi)\|\le\|d_\alpha\phi\|$). Hence, $\phi(a,S^1)\subset\mathbb{C}^n$ is contained in the ball centered at $z_a:=(0,y_a)$ of radius ρ. Similarly, there is some $b\in[1-\delta,1]$ and $y_b\in\mathbb{C}^q$ such that $\phi(b,S^1)\subset\mathbb{C}^n$ is contained in the ball centered at $z_b:=(0,y_b)$ of the same radius ρ.

Let $Z':=[a,b]\times S^1$ and let

$$R:=l_{\min}^{-1}\delta^{-1/2}K(\|d_\alpha\phi\|_{L^2}+1).$$

We claim that $\phi(Z')$ is contained at least in one of the balls $B(z_a,R)$ or $B(z_b,R)$. Suppose this is not the case. Then $\phi(Z')$ is not contained in the union $B(z_a,R/3)\cup B(z_b,R/3)$. Indeed, if this were true, then, since $\phi(Z')$ is connected, we should have $B(z_a,R/3)\cap B(z_b,R/3)\neq 0$ (otherwise $\phi(Z')$ would be included either in $B(z_a,R/2)$ or in $B(z_b,R/2)$, in contradiction with our assumption). But in this case $d(z_a,z_b)<2R/3$, whence $\phi(Z')\subset B(z_a,R/3)\cup B(z_b,R/3)\subset B(z_a,R)$, which we assume not to be true. It follows from all this that there is some $z\in\phi(Z')$ satisfying simultaneously

$$d(z,z_a)>R/3 \qquad\text{and}\qquad d(z,z_b)>R/3.$$

In other words, the ball $B(z,R/6)$ is disjoint both with $B(z_a,R/6)$ and $B(z_b,R/6)$. By our choice of R we have $R\ge 6\rho$, so the discussion above proves that $\phi(a,S^1)\subset B(z_a,R/6)$ and $\phi(b,S^1)\subset B(z_b,R/6)$. We deduce that $B(z,R/6)\cap\phi(\partial Z')=\emptyset$. Hence we can apply Lemma 10.17 to $(\phi,\alpha):Z'\to\mathbb{C}^n$ and deduce that

$$2\pi+\|d_\alpha\phi\|^2_{L^2}\ge \operatorname{area}(Z')+\|d_\alpha\phi\|^2_{L^2(Z')}\ge K''R^2/36.$$

However, it follows from (10.55) that we must have $K''R^2/36>2\pi+\|d_\alpha\phi\|^2_{L^2}$, which is a contradiction. This finishes the proof.

11. Long cylinders with small energy and nearly critical residue

In this section we use the same notations as in the previous one.

For any natural number N, let $C_N=[-N,N]\times S^1$ with the standard product metric and the induced conformal structure. We denote as always by (t,θ) the usual coordinates

in C_N. Let $\upsilon = f dt \wedge d\theta$ be a volume form, and let $\eta > 0$ be any number. We say that υ is **exponentially η-bounded** if

$$|f| < \eta e^{|t|-N}. \tag{11.56}$$

Volume forms satisfying this property arise when we consider C_N as a conformal model for open sets U in stable curves $(C, \mathbf{x})$ of the form $U \simeq \{xy = \delta\} \subset \mathbb{C}^2$ for small δ, and we take on C_N the volume form corresponding to the metric $\nu_{[C,\mathbf{x}]}$.

Theorem 11.1. *Fix some number $c \in \mathbf{i}\mathbb{R}$. There exist numbers $\epsilon > 0$, $\eta_0 > 0$, $\sigma > 0$ and $K > 0$ (depending only on X, the action of S^1 on X, and I) with the following property. Let $(\alpha, \phi) : C_N \times S^1 \to X$ be a pair. Suppose that $\overline{\partial}_{I,\alpha}\phi = 0$ and that there exists some critical residue $\lambda_{\mathrm{cr}} \in \Lambda_{\mathrm{cr}}$ such that*

$$\|\alpha - \lambda_{\mathrm{cr}} d\theta\|_{L^\infty} < \epsilon \qquad \text{and} \qquad \|d_\alpha \phi\|_{L^\infty} < \epsilon. \tag{11.57}$$

Then there exist maps $\psi : [-N, N] \to X^{\lambda_{\mathrm{cr}}}$ and $\phi_0 : C_N \to TX$ satisfying $\phi_0(t, \theta) \in T_{\psi(t)}X$,

$$\phi(t, \theta) = e^{-\lambda_{\mathrm{cr}}\theta} \exp^g_{\psi(t)}(e^{\lambda_{\mathrm{cr}}\theta}\phi_0(t, \theta)) \qquad \text{and} \qquad \int e^{\lambda_{\mathrm{cr}}\theta}\phi_0(t, \theta) d\theta = 0, \tag{11.58}$$

where $\exp^g_x : T_x X \to X$ denotes the exponential map on X with respect to the metric g (see also Remark 11.2 for the precise meaning of the formulae). We distinguish two cases.

(1) *Let υ be an exponentially η-bounded volume form on C_N, where $\eta < \eta_0$. Suppose that*

$$\iota_\upsilon d\alpha + \mu(\phi) = c; \tag{11.59}$$

then the following inequality holds:

$$|\phi_0(t, \theta)| < K e^{-\sigma(N-|t|)} \left(\|d_\alpha \phi\|_{L^\infty} + \|\alpha\|_{L^\infty} + \eta(|c| + \sup|\mu|)\right)^{1/4}. \tag{11.60}$$

(2) *Suppose now that $d\alpha = 0$; then the previous estimate can be improved to:*

$$|\phi_0(t, \theta)| < K e^{-\sigma(N-|t|)}(\|d_\alpha \phi\|_{L^\infty} + \|\alpha\|_{L^\infty})^{1/4}. \tag{11.61}$$

Remark 11.2. *Note that the first formula in (11.58) is well defined. To see this, denote $\lambda_{\mathrm{cr}} = \mathbf{i}p/q$, where p and q are relatively primer integers. Then the condition $\psi \subset X^{\lambda_{\mathrm{cr}}}$ implies that for any t the point $\psi(t)$ is fixed by the action of $e^{2\pi\mathbf{i}/q}$. On the other hand, $e^{\pm\lambda_{\mathrm{cr}}\theta}$ is well defined up to multiplication by powers of $e^{2\pi\mathbf{i}/q}$. Finally, since the exponential map is equivariant we have $e^{-2\pi\mathbf{i}/q}\exp_{\psi(t)}(e^{2\pi\mathbf{i}/q}v) = \exp_{\psi(t)}(v)$ for any vector $v \in T_{\psi(t)}X$. As for the second formula in (11.58), it should be understood as*

$$\int e^{\lambda_{\mathrm{cr}}\theta}\phi(t, \theta) d\theta = \int_0^{2\pi q} e^{\lambda_{\mathrm{cr}}\theta}\phi(t, \theta) d\theta.$$

The proof of Theorem 11.1 will be given in Section 11.1 below. Next theorem states that the map ψ constructed in Theorem 11.1 in approximately a gradient flow line for the moment map μ. Before stating it we introduce some notation. Suppose that (α, ϕ) is a pair satisfying the hypothesis of Theorem 11.1. Applying a gauge transformation if necessary, we can assume that that α is in temporal gauge, $\alpha = \alpha_\theta d\theta$, and that α_θ restricted to $\{0\} \times S^1$ takes a constant value $\lambda + \lambda_{\mathrm{cr}} \in \mathbf{i}\mathbb{R}$. Then we can write

$$\alpha = (\lambda + \lambda_{\mathrm{cr}} + \beta) d\theta \tag{11.62}$$

for some function $\beta : C_N \to \mathbf{i}\mathbb{R}$.

Theorem 11.3. *Following the notation of Theorem 11.1, we distinguish again two cases.*

(1) *Suppose that $\iota_v d\alpha + \mu(\phi) = c$, where v is η-bounded and $\eta < \eta_0$; then*

$$|\psi'(t) + \mathbf{i}\lambda I(\psi(t))\mathcal{X}(\psi(t))| < K e^{-\sigma(N-|t|)} \left(\|d_\alpha \phi\|_{L^\infty} + \|\alpha\|_{L^\infty} + \eta(|c| + \sup|\mu|) \right)^{1/4}, \tag{11.63}$$

holds for any t, where K is a constant depending only on X, the action of S^1, and the almost complex structure I;

(2) *suppose that $d\alpha = 0$; then, for any t, we have*

$$|\psi'(t) + \mathbf{i}\lambda I(\psi(t))\mathcal{X}(\psi(t))| < K e^{-\sigma(N-|t|)} (\|d_\alpha \phi\|_{L^\infty} + \|\alpha\|_{L^\infty})^{1/4}, \tag{11.64}$$

The proof of Theorem 11.3 will be given in Section 11.2

11.1. **Proof of Theorem 11.1.** We first prove the theorem under the assumption that the critical residue λ_{cr} is equal to 0.

The existence of ψ and ϕ_0 satisfying (11.58) follows easily from the implicit function theorem (see for example Section 14 in [FO]) provided ϵ is small enough (so that $\|\alpha\|_{L^\infty} < \epsilon$ together with $\|d_\alpha \phi\|_{L^2} < \epsilon$ imply that for any t the image of $\phi(t, \cdot) : S^1 \to X$ is contained in a small ball of radius less than the injectivity radius of X and g).

To prove that the estimate (11.60) follows from (11.59) we use the same strategy as in the proof of Theorem 10.4, namely, we make use of an inequality involving the L^2 norm of ϕ_0 restricted to three consecutive pieces of the cylinder and then using the fact that (α, ϕ) is I-holomorphic to deduce pointwise bounds on ϕ_0 from L^2 bounds. To state the inequality we define as in the previous section $C = [-2, 3] \times S^1$. Recall also that when we write any norm (L^∞, L^2, etc.) of either the connection or the section of a pair defined over C, or of a complex structure depending on points of C, we mean the norm over C (unless we specify some other domain). Let also Z_I, Z_{II} and Z_{III} be the subsets of C defined in (10.35).

Theorem 11.4. *There exist some constants $\epsilon > 0$ and $K > 0$, depending only on X and I, with the following property. Let $(\alpha, \phi) : C \times S^1 \to X$ be a I-holomorphic pair satisfying $\|\alpha\|_{L^\infty} < \epsilon$ and $\|d_\alpha \phi\|_{L^\infty} < \epsilon$. Let $\psi : [-2, 3] \to X$ and $\phi_0 : C \to TX$ be defined as in Theorem 11.1.*

(1) *Suppose that* $\|d\alpha\|_{L^2(C)} < \epsilon\|\phi_0\|_{L^2(C)}$. *Then*

$$\|\phi_0\|^2_{L^2(Z_{II})} \leq \frac{1}{e+e^{-1}}\left(\|\phi_0\|^2_{L^2(Z_I)} + \|\phi_0\|^2_{L^2(Z_{III})}\right) \tag{11.65}$$

(2) *Suppose now that* $\|d\alpha\|_{L^2_1} < K$ *and* $\|d\alpha\|_{L^\infty} < K$. *Then* $\sup_Z |\phi_0| \leq K\|\phi_0\|^{1/4}_{L^2}$.

Theorem 11.4 will be proved in Section 11.3, and we now prove Theorem 11.1. Take ϵ as in Theorem 11.4. Since $d_\alpha\phi = d\phi - \mathbf{i}\alpha\mathcal{X}(\phi)$, we can bound

$$\|d\phi\|_{L^\infty} \leq K(\|d_\alpha\phi\|_{L^\infty} + \|\alpha\|_{L^\infty}). \tag{11.66}$$

As a consequence, and taking (11.57) into account, we deduce that $\|d\phi\|_{L^\infty}$ is uniformly bounded. Let $v = f\,dt \wedge d\theta$. Equation (11.59) can be written as $d\alpha = f(c - \mu(\phi))dt \wedge d\theta$. Since $\mu(\phi)$ is uniformly bounded (because X is compact), once c has been chosen we can take η_0 in such a way that the requirement of v being exponentially η-bounded (and hence η_0-bounded as well) implies that $\|d\alpha\|_{L^2_1(C_N)}$ is necessarily less than the K in Theorem 11.4 (here we use that $\|d\phi\|_{L^\infty}$ is uniformly bounded). On the other hand, we have

$$|d\alpha|(t,\theta) < \eta(|c| + \sup|\mu|)e^{|t|-N}. \tag{11.67}$$

In the sequel we will denote for convenience $R := \eta(|c| + \sup|\mu|)$.

Define, for every $-N \leq n \leq N$, $Z_n = [n, n+1]$ and $x_n := \|\phi_0\|_{L^2(Z_n)}$. Formula (11.66) implies that for any n we have

$$x_n \leq K(\|d_\alpha\phi\|_{L^\infty} + \|\alpha\|_{L^\infty}). \tag{11.68}$$

Let also $z_n := \|d\alpha\|_{L^2(Z_n)}$. Formula (11.67) implies that $z_n \leq K_0 R e^{-(N-|n|)}$ for every n, where K_0 is a universal constant. On the other hand, Theorem 11.4 implies that if

$$z_{n-2} + \cdots + z_{n+2} \leq \epsilon(x_{n-2} + \cdots + x_{n+2})$$

then $x_n \leq (e + e^{-1})^{-1}(x_{n-1} + x_{n+1})$. Hence the sequences $\{x_n\}$ and $\{z_n\}$ satisfy the requirements of Lemma 10.9. Applying Lemma 10.9 together with (11.68) we deduce, for some $\sigma_0 > 0$ independent of (α, ϕ), an estimate

$$\|\phi_0\|_{L^2(Z_j)} = x_j \leq \epsilon^{-1}K_0(R + \|d_\alpha\phi\|_{L^\infty} + \|\alpha\|_{L^\infty})e^{-\sigma_0(N-|j|)}.$$

To finish the proof of the theorem and obtain (11.60), combine the previous inequality with (2) of Theorem 11.4 (the two necessary conditions for section (2) of Theorem 11.4 are satisfied in our situation: the L^2_1 norm of $d\alpha$ on the whole C_N is less than K and $d\alpha$ has bounded L^∞ norm by (11.67)).

The proof that $d\alpha = 0$ implies (11.61) follows exactly the same scheme (and is even easier).

Now we consider the case of general critical residue $\lambda_{\mathrm{cr}} \in \Lambda_{\mathrm{cr}}$. Let us write $\lambda_{\mathrm{cr}} = \mathbf{i}p/q$ for some relatively prime integers p, q satisfying $q \geq 1$. Consider the covering map

$$\pi : C_{N/q} \to C_N$$

defined as $\pi(t,\theta) := (qt, q\theta)$, and define a new pair $(\alpha', \phi') := \pi^*(\alpha, \phi)$. The first inequality in (11.57) is equivalent to $\|\alpha' - p\,d\theta\|_{L^\infty} < q\epsilon$. Applying the gauge transformation

$g(t,\theta) := e^{\mathbf{i}p\theta}$ we obtain another pair $(\alpha'',\phi'') := g^*(\alpha',\phi') = (\alpha' - pd\theta, e^{\mathbf{i}\theta p}\phi')$. The new connection α'' now satisfies $\|\alpha''\|_{L^\infty} < q\epsilon$, $\|d_{\alpha''}\phi''\|_{L^\infty} < q\epsilon$ and $\overline{\partial}_{I,\alpha''}\phi'' = 0$. Also, equation (11.59) is satisfied with v replaced by $\pi^* v$ (which now is exponentially $q\eta$-bounded). Hence, provided ϵ and η are small enough, we can apply the case of zero residue to the pair (α'',ϕ''). We thus arrive at a pair of maps $\psi'' : [-N/q, N/q] \to X$ and $\phi''_0 : C_{N/q} \to TX$ satisfying $\phi'' = \exp_{\psi''}\phi''_0$ and $\int \phi''_0(t,\nu)d\nu = 0$. We now prove that $\psi'' \subset X^{\lambda_{\mathrm{cr}}}$. Let $\alpha := 2\pi/q$. For any t,θ we compute

$$\phi''(t,\alpha+\theta) = e^{\mathbf{i}(\alpha+\theta)p}\phi'(t,\alpha+\theta) = e^{\mathbf{i}\alpha p}e^{\mathbf{i}\theta p}\phi'(t,\theta) = e^{\mathbf{i}\alpha p}\phi''(t,\theta),$$

which, combined with the fact that g is S^1-invariant, implies $\psi'' \subset X^{\lambda_{\mathrm{cr}}}$. The computation above implies similarly that $\phi''_0(t,\alpha+\theta) = e^{\mathbf{i}\alpha p}\phi''_0(t,\theta)$, so that the map

$$\widetilde{\phi}''_0(t,\theta) := e^{-\mathbf{i}\theta p}\phi''_0(t,\theta)$$

satisfies $\widetilde{\phi}''_0(t,\theta) = \widetilde{\phi}''_0(t,\alpha+\theta)$ and hence descends to give a map $\phi_0 : C_N \to X$ such that

$$\widetilde{\phi}''_0(t,\theta) = \phi_0(qt,q\theta).$$

Define also $\psi : [-N,N] \to X^{\lambda_{\mathrm{cr}}}$ by the condition $\psi''(t) = \psi(qt)$. Then the equality $\phi'' = \exp_{\psi''}\phi''_0$ translates into

$$\phi(qt,q\theta) = e^{-\mathbf{i}\theta p}\exp_{\psi(qt)}(e^{\mathbf{i}\theta p}\phi_0(qt,q\theta)),$$

which is equivalent to the first formula in (11.58). On the other hand, the balancing condition $\int \phi''_0(t,\nu)d\nu = 0$ is clearly equivalent to $\int_0^{2\pi q} e^{\lambda_{\mathrm{cr}}\nu}\phi_0(t,\nu)d\nu = 0$.

Finally, since the set of q which appear as denominators of numbers in $\mathbf{i}\Lambda_{\mathrm{cr}}$ is finite, ϵ and η_0 can be chosen in such a way that for any critical residue we can reduce to the case of zero residue applying the procedure which we just described.

11.2. **Proof of Theorem 11.3.** As in the proof of Theorem 11.1, we prove the case $\lambda_{\mathrm{cr}} = 0$. Using the same covering argument as in Section 11.1 the general case is reduced to this case. Furthermore, we only prove that $\iota_v d\alpha + \mu(\phi) = c$ implies (11.63). That $d\alpha = 0$ implies (11.64) is even simpler and follows from the same ideas.

We begin by stating a local version of the theorem.

Lemma 11.5. *Let $K_0 > 0$ be a real number, let I (resp. g, $\mathcal{X}$) be an almost complex structure (resp. Riemannian metric, vector field) on $\mathbb{C}^n$, and let $V \subset \mathbb{C}^n$ be a compact subset. Let $Z := [-1,2] \times S^1$, and let $\phi : Z \to V$ be a map satisfying*

$$\frac{\partial\phi}{\partial t} = I(\phi)\left(\frac{\partial\phi}{\partial\theta} - \mathbf{i}(\lambda+\beta)\mathcal{X}(\phi)\right). \tag{11.69}$$

Suppose that the diameter of $\phi(Z)$ is small enough so that $\psi : [0,1] \to \mathbb{C}^n$ and $\phi_0 : Z \to \mathbb{C}^n$ can be defined as in Theorem 11.1 (here we identify $T_{\psi(t)}\mathbb{C}^n \simeq \mathbb{C}^n$) and that

$$|\phi_0| < K_0 \qquad \textit{and} \qquad \left|\frac{\partial\phi_0}{\partial\theta}\right|, \left|\frac{\partial\phi_0}{\partial t}\right| < K_0. \tag{11.70}$$

Then for any $t \in (0,1)$ *we have*

$$|\psi'(t) + \mathbf{i}\lambda I(\psi(t))\mathcal{X}(\psi(t))| < K \sup_{Z_t}(|\phi_0| + |\beta|) \tag{11.71}$$

for some constant K *independent of* ϕ*, where* $Z_t := \{t\} \times S^1$*.*

Proof. Let $\widetilde{V} := \{\exp^g_x v \mid x \in V, |v| \le K_0\}$. By hypothesis we have $\psi([0,1]) \subset \widetilde{V}$. Let $B \subset \mathbb{C}^n$ be the closed ball of radius K_0, let $E : \widetilde{V} \times B \to \mathbb{C}^n$ be the exponential map $E(x,v) := \exp^g_x v$. Consider the first derivatives $F := D_x E$ and $G := D_v E$. Since the domain of E is compact, the second derivatives of E are uniformly bounded; combining this observation with the fact that $F(x,0) = \mathrm{Id}$ and $G(x,0) = \mathrm{Id}$ for any x, we obtain bounds

$$|F(x,v) - \mathrm{Id}\,| < K|v| \qquad \text{and} \qquad |G(x,v) - \mathrm{Id}\,| < K|v|. \tag{11.72}$$

Taking partial derivatives with respect to t in the equality $\exp_\psi \phi_0 = \phi$ we obtain

$$F(\psi,\phi_0)\psi' + G(\psi,\phi_0)\frac{\partial \phi_0}{\partial t} = \frac{\partial \phi}{\partial t}.$$

Writing $F = \mathrm{Id} + (F - \mathrm{Id})$ and $G = \mathrm{Id} + (G - \mathrm{Id})$ and integrating over S^1 the above equality, combined with (11.70), (11.72), and the fact that $\int \phi_0(t,\theta)d\theta = 0$ for every t, yields

$$\left|\psi'(t) - \frac{1}{2\pi}\int \frac{\partial \phi}{\partial t}\right| < K \sup_{Z_t} |\phi_0|. \tag{11.73}$$

Using (11.69) we compute

$$\begin{aligned}\frac{\partial \phi}{\partial t} =& - \mathbf{i}\lambda I(\psi)\mathcal{X}(\psi) + I(\psi)\frac{\partial \phi}{\partial \theta} + (I(\phi) - I(\psi))\left(\frac{\partial \phi}{\partial \theta} - \mathbf{i}\lambda \mathcal{X}(\phi)\right) \\ &- \mathbf{i}\lambda I(\psi)(\mathcal{X}(\phi) - \mathcal{X}(\psi)) - \mathbf{i}\beta I(\phi)\mathcal{X}(\phi).\end{aligned}$$

Integrating over S^1 and dividing by 2π the first term in the right hand side does not change; the second term vanishes, because $I(\psi)$ is independent of θ and the integral of $\partial\phi/\partial\theta$ over S^1 is obviously 0; the third and fourth term can be bounded by $K \sup_{Z_t} |\phi_0|$, since both $|I(\psi) - I(\phi)|$ and $|\mathcal{X}(\psi) - \mathcal{X}(\phi)|$ are less than $K \sup_{Z_t} |\phi_0|$ (of course here K depends on the derivatives of I and $\mathcal{X}$) and (11.70) holds; the fifth term can be bounded by $K \sup_{Z_t} |\beta|$. Putting together all this observations and using (11.73) we obtain inequality (11.71). □

Covering X with a finite number of charts we deduce from the lemma that for any $t \in [-N,N]$:

$$|\psi'(t) + \mathbf{i}\lambda I(\psi(t))\mathcal{X}(\psi(t))| < K \sup_{\{t\}\times S^1} (|\phi_0| + |\beta|). \tag{11.74}$$

Note in fact that the constant K_0 in the Lemma can be taken independently of (α,ϕ), and depending only on the ϵ in Theorem 11.1, because (11.57) implies the bounds (11.70). Combining this with the fact that the number of charts which we use is finite, we deduce that the constant K in the above inequality is independent also of (α,ϕ).

To finish the proof of Theorem 11.3 we need to bound the right hand side of (11.74). On the one hand, formula (11.60) in Theorem 11.1 tells us that

$$|\phi_0(t,\theta)| < Ke^{\sigma(|t|-N)} \left(\|d_\alpha\phi\|_{L^\infty} + \|\alpha\|_{L^\infty} + \eta(|c| + \sup|\mu|)\right)^{1/4}. \tag{11.75}$$

On the other hand, we can bound β writing

$$\beta(t,\theta) = \beta(0,\theta) + \int_0^t \frac{\partial\beta}{\partial t}(\tau,\theta)d\tau = \int_0^t \frac{\partial\alpha_\theta}{\partial t}(\tau,\theta)d\tau,$$

then using $d\alpha = \frac{\partial\alpha_\theta}{\partial t}dt \wedge d\theta$ and (11.67) to obtain

$$\begin{aligned} |\beta(t,\theta)| &\le \int_0^t |d\alpha(\tau,\theta)|d\tau \le \int_0^t e^{|\tau|-N}\eta(|c| + \sup|\mu|)d\tau \\ &\le e^{|\tau|-N}\eta(|c| + \sup|\mu|) \le e^{\sigma(|t|-N)}\eta(|c| + \sup|\mu|) \\ &\le Ke^{\sigma(|t|-N)}(\eta(|c| + \sup|\mu|))^{1/4} \end{aligned} \tag{11.76}$$

where we have used that $\sigma < 1$ and the fact that $\eta(|c| + \sup|\mu|)$ is bounded above. Combining the estimates (11.75) with (11.76) with (11.74) we obtain the desired bound.

11.3. **Proof of Theorem 11.4.** Using Lemma 10.4 we reduce the proof of Theorem 11.4 to a local version of it, very much as we did in the proof of Theorem 10.5. We use the same notation as in Lemma 10.6, so we fix a diagonal action of S^1 on $\mathbb{C}^n$ with weights $w = (w_1, \dots, w_n) \in \mathbb{Z}^n$. Let I_0 (resp. g_0) denote the standard complex structure (resp. Riemannian metric) on $\mathbb{C}^n$.

Lemma 11.6. *For any $r > 0$ and $K > 0$ there exists some $\epsilon = \epsilon(K, w, r) > 0$ with the following property. Suppose that I (resp. g) is a smooth equivariant almost complex structure (resp. Riemannian metric) on $\mathbb{C}^n$ such that $\|I - I_0\|_{L^\infty} < \epsilon$, $\|DI\|_{L^\infty} < K$, and $\|D^2I\|_{L^\infty} < K$ (here DI and D^2I denote the first and second derivatives of I and the norms are taken with respect to the standard metric g_0 on $\mathbb{C}^n$), and $\|g - g_0\|_{L^\infty} < \epsilon$ and $\|Dg\|_{L^\infty} < K$. Let $(\alpha, \phi) : C \times S^1 \to \mathbb{C}^n$ be a I-holomorphic pair satisfying $\|\alpha\|_{L^\infty} < \epsilon$, $\|d_\alpha\phi\|_{L^\infty} < \epsilon$ and*

$$\phi(C) \subset B(0, 2r) \subset \mathbb{C}^n.$$

Let $\psi : [-2, 3] \to \mathbb{C}^n$ and $\phi_0 : C \to T\mathbb{C}^n$ be defined as in Theorem 11.1.

(1) *Suppose that $\|d\alpha\|_{L^2} < \epsilon\|\phi_0\|_{L^2}$. Then*

$$\|\phi_0\|^2_{L^2(Z_{II})} \le \frac{1}{e + e^{-1}} \left(\|\phi_0\|^2_{L^2(Z_I)} + \|\phi_0\|^2_{L^2(Z_{III})}\right) \tag{11.77}$$

(2) *Suppose now that $\|d\alpha\|_{L^2_1} < K$ and $\|d\alpha\|_{L^\infty} < K$. Then $\sup_Z |\phi_0| \le K\|\phi_0\|^{1/4}_{L^2}$.*

The proof of the lemma will be given in Section 11.5 below.

Remark 11.7. *We specify the dependence of ϵ on DI, D^2I and Dg because we want to be sure that the hypothesis of the lemma are preserved when we zoom in (this will become clear in the course of the proof). The same comment applies for Lemma (11.8).*

11.4. **Comparison between ϕ_0 and ϕ_{av}.** When (α, ϕ) is a pair taking values in $\mathbb{C}^n$, the map ϕ_0 defined in Theorem 11.4 can be roughly speaking approximated by the map ϕ_{av} defined in (2) of Lemma 10.10. The following lemma makes this statement precise, specifying to what extent ϕ_{av} is a good approximation of ϕ_0, both in L^2 and C^0 norms. This will be crucial in proving Lemma 11.6, since it is clearly easier to deal with ϕ_{av} than with ϕ_0.

Lemma 11.8. *For any $K > 0$ there exist constants $K_0 > 0$ and $\epsilon > 0$ with the following property. Denote by g_0 the standard Riemannian flat metric in $\mathbb{C}^n$. Let g be another metric on $\mathbb{C}^n$ such that $K^{-1}g_0 \leq g \leq Kg_0$ and such that $\|Dg\|_{L^\infty} < K$, where Dg denotes the first derivatives of g and the norm is taken with respect to g_0. Let $x \in M$ and let $\gamma_0 : S^1 \to \mathbb{C}^n$ be a smooth map satisfying $\int \gamma_0 = 0$ and $\sup|\gamma_0| < \epsilon$. Let $\gamma(\theta) := \exp^g_x \gamma_0(\theta)$ (here we are identifying $T_x\mathbb{C}^n \simeq \mathbb{C}^n$) and define*

$$\gamma_{\mathrm{av}}(\theta) := \gamma(\theta) - \frac{1}{2\pi}\int \gamma(\nu)d\nu.$$

Then we have

$$\|\gamma_0\|_{L^2}(1 - K_0\|\gamma_0\|_{L^2}) \leq \|\gamma_{\mathrm{av}}\|_{L^2} \leq \|\gamma_0\|_{L^2}(1 + K_0\|\gamma_0\|_{L^2}) \tag{11.78}$$

and similarly

$$\sup|\gamma_0|(1 - K_0 \sup|\gamma_0|) \leq \sup|\gamma_{\mathrm{av}}| \leq \sup|\gamma_0|(1 + K_0 \sup|\gamma_0|). \tag{11.79}$$

Proof. We claim that given $K > 0$, $\epsilon > 0$, and a metric g on $\mathbb{C}^n$ satisfying $K^{-1}g_0 \leq g \leq Kg_0$ and such that $\|Dg\|_{L^\infty} < K$, there is a constant K' depending only on K such that for any $x, v \in \mathbb{C}^n$ satisfying $|v| < \epsilon$ we have

$$|\exp^g_x v - x - v| < K'|v|^2. \tag{11.80}$$

To see this, define $\gamma(t) := \exp^g_x tv$. Since $\gamma(t)$ is a geodesic, we have $\frac{d\gamma_k'}{dt} = -\Gamma^k_{ij}\gamma_i'\gamma_j'$ (here $\gamma_1, \ldots, \gamma_{2n}$ are the components of γ). Integrating and using $\gamma'(0) = v$ we deduce that (provided $\epsilon > 0$ has been chosen small enough and $|v| < \epsilon$) for any $t \in [0,1]$, $|\gamma'(t) - v| < K'|v|^2$, where K' is proportional to the sup norm of Γ^k_{ij} which, on its turn, can be estimated in terms of $(g^{ij}) = (g_{ij})^{-1}$ and the derivatives of (g_{ij}). Integrating this inequality we obtain (11.80).

It follows from that (11.80) that

$$\left|x - \frac{1}{2\pi}\int \gamma\right| = \left|x - \frac{1}{2\pi}\int \exp^g_x \gamma_0\right| < K'\|\gamma_0\|^2_{L^2}.$$

Consequently we have

$$\begin{aligned}
|\gamma_{\mathrm{av}}(\theta) - \gamma_0(\theta)| &= \left|\gamma(\theta) - \left(\frac{1}{2\pi}\int\gamma\right) - \gamma_0(\theta)\right| \\
&< \left|x + \gamma_0(\theta) - \left(\frac{1}{2\pi}\int\gamma\right) - \gamma_0(\theta)\right| + K'|\gamma_0(\theta)|^2 \\
&< K'(\|\gamma_0\|^2_{L^2} + |\gamma_0(\theta)|^2).
\end{aligned}$$

Integrating over θ we obtain (11.78). Inequality (11.79) is proved similarly. □

11.5. **Proof of Lemma 11.6.** We follow an idea similar to the proof of Lemma 10.6 in Section 10.8. Assume that there exists sequences of positive real numbers $\epsilon_u \to 0$, invariant almost complex structures I_u and metrics g_u, and I_u-holomorphic pairs (α_u, ϕ_u). Suppose that I_u, g_u and (α_u, ϕ_u) satisfy the hypothesis of the lemma for $\epsilon = \epsilon_u$. In particular, we have

$$\|\alpha_u\|_{L^\infty} \to 0 \qquad \text{and} \qquad \|d_{\alpha_u}\phi_u\|_{L^\infty} \to 0. \tag{11.81}$$

Finally, suppose that for each u either (1) or (2) of the lemma fails to be true. We will see that this is impossible.

As we did in Lemma 10.6, we assume that w_j is nonzero for any j between 1 and p, and that $w_{p+1} = \cdots = w_n = 0$. Let $q = n - p$, so that we have a splitting $\mathbb{C}^n = \mathbb{C}^p \times \mathbb{C}^q$. Let also π_p and π_q denote the projections from $\mathbb{C}^n$ to $\mathbb{C}^p$ and $\mathbb{C}^q$.

Define $\rho_u := \|d_{\alpha_u}\phi_u\|_{L^\infty}$ and $s_u := \sup|\pi_p\phi_u(C)|$. Note that since $\phi_u(C) \subset B(0, 2r)$ the numbers s_u are uniformly bounded above. Let $y_u \in C$ be a point where $|\pi_p\phi_u(C)|$ attains the value s_u, and let $x_u = \phi_u(y_u)$. Passing to a subsequence if necessary, we can assume that the sequence $\rho_u^{-1}s_u$ converges somewhere in $\mathbb{R}_{\geq 0} \cup \{\infty\}$. We distinguish two possibilities.

Suppose first that $\rho_u^{-1}s_u \to e < \infty$. Let $\phi'_u := \rho_u^{-1}(\phi_u - x_u)$. Define also I'_u (resp. g_u) to be the pullback of I_u (resp. g_u) under the composition of the translation along x_u with the homotecy of ratio ρ. Then I'_u and g'_u are S^1-invariant and still satisfy the bounds in the hypothesis of the lemma. Furthermore, (α_u, ϕ'_u) is a I'_u-holomorphic pair. Finally, since $\|d_{\alpha_u}\phi'_u\|_{L^\infty} = 1$, $\|\alpha\|_{L^\infty} \to 0$ and $|\phi'_u(y)| \leq e$, the image $\phi'_u(C)$ is contained in a compact set independent of u. Hence we may apply Lemma 10.13 and deduce that, up to regauging, the pairs (α_u, ϕ'_u) converge in C^1 norm to a I_0-holomorphic pair $(0, \phi)$. So $\|d\phi\|_{L^\infty} = 1$ (hence ϕ is not constant) and $\overline{\partial}_{I_0}\phi = 0$.

Let ϕ_{av} be as defined in (2) of Lemma 10.10. Since ϕ is holomorphic and not constant, by Lemma 10.10 we have

$$\|\phi_{\mathrm{av}}\|_{L^2(Z_{II})} \leq \frac{1}{e^2 + e^{-2}}\|\phi_{\mathrm{av}}\|_{L^2(Z_I)} + \|\phi_{\mathrm{av}}\|_{L^2(Z_{III})}).$$

On the other hand, since the convergence $\phi'_u \to \phi$ is in C^0, it follows that $\phi'_{u,\mathrm{av}}$ converges pointwise to ϕ_{av}. This implies that for big enough u we have

$$\|\phi'_{u,\mathrm{av}}\|_{L^2(Z_{II})} < \frac{1}{e^{1.5} + e^{-1.5}}(\|\phi'_{u,\mathrm{av}}\|_{L^2(Z_I)} + \|\phi'_{u,\mathrm{av}}\|_{L^2(Z_{III})}).$$

Since ϕ'_u is related to ϕ_u by a translation and a homotecy, it follows that the same inequality is satisfied by ϕ_u. Finally, (11.78) in Lemma 11.8 implies (using the fact that $\|\phi_{u,0}\|_{L^2}$ converges to 0) that for big enough u

$$\|\phi_{u,0}\|_{L^2(Z_{II})} < \frac{1}{e + e^{-1}}(\|\phi_{u,0}\|_{L^2(Z_I)} + \|\phi_{u,0}\|_{L^2(Z_{III})}).$$

Hence (1) in Lemma 11.6 has to hold for big enough u.

On the other hand, since ϕ_{av} is holomorphic and satisfies $\int \phi_{\mathrm{av}}(t,\nu)d\nu = 0$, standard elliptic estimates imply that, for some K independent of ϕ, we have

$$\sup_Z |\phi_{\mathrm{av}}| \leq K\|\phi_{\mathrm{av}}\|_{L^2}.$$

Again using the fact that $\phi'_{u,\mathrm{av}}$ converges to ϕ_{av} we deduce that the same inequality holds for $\phi_{u,\mathrm{av}}$ provided u is big enough and maybe after increasing slightly K. Finally, (11.79) in Lemma 11.8 implies a similar inequality $\sup_Z |\phi_{u,0}| \leq K\|\phi_{u,0}\|_{L^2}$. But since $\|\phi_{u,0}\|_{L^2}$ is smaller than 1 for big enough u, this implies

$$\sup_Z |\phi_{u,0}| \leq K\|\phi_{u,0}\|_{L^2}^{1/4}.$$

Consequently, (2) in Lemma 11.6 has to hold for big enough u. And this is in contradiction with our assumptions.

Now suppose that $\rho_u^{-1}s_u \to \infty$ and define $\phi'_u := s_u^{-1}(\phi_u - x_u)$. Then $\|d_{\alpha_u}\phi'_u\|_{L^\infty} \to 0$, so the diameter of ϕ'_u converges to 0. On the other hand each ϕ'_u intersects the set

$$S = \{(x,0) \in \mathbb{C}^p \times \mathbb{C}^q \mid |x| = 1\} \subset \mathbb{C}^n.$$

Then (1) in Lemma 11.9 below implies that for big enough u the pairs have to satisfy (1) in Lemma 11.6. On the other hand, (2) in Lemma 11.9 implies that for big enough u we have $\sup_Z |\phi'_{u,0}| \leq K\|\phi'_{u,0}\|_{L^2}^{1/4}$. Since $\phi_{u,0} = s_u\phi'_{u,0}$ and s_u is uniformly bounded above, maybe after increasing K we also have the following inequality for big enough u:

$$\sup_Z |\phi_{u,0}| \leq K\|\phi_{u,0}\|_{L^2}^{1/4}.$$

This implies that (2) in Lemma 11.6 has to hold for big enough u, leading to a contradiction and finishing the proof of the lemma.

Lemma 11.9. *For any $K > 0$ there exist numbers $\epsilon = \epsilon(K,w) > 0$ and $\delta = \delta(K,w) > 0$ with the following property. Suppose that I (resp. g) is a smooth equivariant almost complex structure (resp. Riemannian metric) on $\mathbb{C}^n$ such that $\|I - I_0\|_{L^\infty} < \epsilon$, $\|DI\|_{L^\infty} < K$, and $\|D^2I\|_{L^\infty} < K$ and $\|g - g_0\|_{L^\infty} < \epsilon$ and $\|Dg\|_{L^\infty} < K$. Let $(\alpha,\phi) : C \times S^1 \to \mathbb{C}^n$ be a I-holomorphic pair satisfying $\|\alpha\|_{L^\infty} < \epsilon$, $\|d_\alpha\phi\|_{L^\infty} < \epsilon$, and*

$$\phi(C) \subset B(y,\delta) \subset \mathbb{C}^n,$$

where y is a point in S. Let $\psi : [-2,3] \to \mathbb{C}^n$ and $\phi_0 : C \to T\mathbb{C}^n$ be defined as in Theorem 11.1.

(1) *Suppose that $\|d\alpha\|_{L^2} < \epsilon\|\phi_0\|_{L^2}$. Then*

$$\|\phi_0\|_{L^2(Z_{II})}^2 \leq \frac{1}{e + e^{-1}}\left(\|\phi_0\|_{L^2(Z_I)}^2 + \|\phi_0\|_{L^2(Z_{III})}^2\right).$$

(2) *There is a constant K' depending only on K with the following property. Suppose that $\|d\alpha\|_{L_1^2} < K$ and $\|d\alpha\|_{L^\infty} < K$. Then $\sup_Z |\phi_0| \leq K'\|\phi_0\|_{L^2}^{1/4}$.*

11.6. **Proof of Lemma 11.9.** We deduce the lemma from a local statement similar to Lemma 11.6 but accounting for the degenerate situation in which instead of a linear action of S^1 on a vector space we have an action of $\mathbb{R}$ by translations. So now $\mathcal{X}$ denotes a constant vector in $\mathbb{C}^n$.

If $(\alpha, \phi) : C \to \mathbb{C}^n$ is a pair in which $\alpha = \alpha_\theta d\theta$, its energy density is defined by

$$|d_\alpha \phi|^2 = \left|\frac{\partial \phi}{\partial t}\right|^2 + \left|\frac{\partial \phi}{\partial \theta} - \mathbf{i}\alpha_\theta \mathcal{X}\right|^2.$$

Furthermore, the condition on (α, ϕ) of being I-holomorphic reads

$$\frac{\partial \phi}{\partial t} = I(\phi)\left(\frac{\partial \phi}{\partial \theta} - \mathbf{i}\alpha_\theta \mathcal{X}\right). \tag{11.82}$$

Lemma 11.10. *For any $K > 0$ there exists a constant $\epsilon > 0$ with the following property. Let I be a translation invariant almost complex structure on $\mathbb{C}^n$ satisfying*

$$\|I - I_0\|_{L^\infty} < \epsilon, \qquad \|DI\|_{L^\infty} < K, \qquad \text{and} \qquad \|D^2 I\|_{L^\infty} < K$$

and let g be a Riemannian metric on $\mathbb{C}^n$ such that $\|g - g_0\|_{L^\infty} < \epsilon$ and $\|Dg\|_{L^\infty} < K$. Take a I-holomorphic pair $(\alpha, \phi) : C \to \mathbb{C}^n$ satisfying

$$\|\alpha\|_{L^\infty} < \epsilon, \qquad \text{and} \qquad \|d_\alpha \phi\|_{L^\infty} < \epsilon, \tag{11.83}$$

and define $\psi : [-2, 3] \to \mathbb{C}^n$ and $\phi_0 : C \to T\mathbb{C}^n$ as in Theorem 11.1.

(1) *Suppose that $\|d\alpha\|_{L^2} < \epsilon \|\phi_0\|_{L^2}$. Then*

$$\|\phi_0\|^2_{L^2(Z_{II})} \le \frac{1}{e + e^{-1}}\left(\|\phi_0\|^2_{L^2(Z_I)} + \|\phi_0\|^2_{L^2(Z_{III})}\right). \tag{11.84}$$

(2) *There is a constant K' depending only on K with the following property. Suppose that $\|d\alpha\|_{L^2_1} < K$ and $\|d\alpha\|_{L^\infty} < K$. Then we have*

$$\sup_Z |\phi_0| \le K' \|\phi_0\|^{1/4}_{L^2}. \tag{11.85}$$

We now resume the proof of Lemma 11.9. To avoid confusion, let E denote $\mathbb{C}^n$ with the linear action of S^1 and let F denote $\mathbb{C}^n$ with an action of $\mathbb{R}$ given by translations along a vector $\mathcal{X} \in F$. We have the following analogue of the charts constructed in Lemma 10.4. Suppose that I (resp. g) is any S^1-invariant complex structure (resp. Riemannian metric) on E. Let $\epsilon > 0$ be as in Lemma 11.10. For any point $y \in S \subset E$ there exists a neighbourhood U_x of y, a translation invariant complex structure I_x and Riemannian metric g_x on F satisfying $\|I_x - I_0\|_{L^\infty} < \epsilon$ and $\|g_x - g_0\|_{L^\infty} < \epsilon$ (where I_0 and g_0 are the standard structures on F) and such that $\|DI\|_{L^\infty}$, $\|D^2 I\|_{L^\infty}$ and $\|Dg\|_{L^\infty}$ are bounded, and a complex isometry $\xi_x : U_x \to F$. Then $\{U_x\}_{x \in S}$ cover S. Consider a finite subcovering $\{U_1, \dots, U_l\}$. For each open set U_j, let

$$K_j := \max\{\|DI_j\|_{L^\infty},\ \|D^2 I_j\|_{L^\infty},\ \|Dg_j\|_{L^\infty}\}.$$

Let ϵ_j be the value of ϵ given by Lemma 11.10 for the choice $K = K_j$, and set $\epsilon := \min\{\epsilon_1, \dots, \epsilon_l\}$. There exists some constant $\delta > 0$ such that any ball of radius δ centered somewhere in S is contained in one of the U_j's. Taking this value of δ, the statement of

Lemma 11.9 follows from considering one of the open sets (say U_j) which contains the image of ϕ, taking the composition $\phi_j := \xi_j \circ \phi$ and applying Lemma 11.10 to (α, ϕ_j).

11.7. **Proof of Lemma 11.10.**

11.7.1. *Proof of (1).* We proceed by contradiction. Assume that there exist sequences of real numbers $\epsilon_u > 0$, almost complex structures I_u, Riemannian metrics g_u on $\mathbb{C}^n$ and I_u-holomorphic pairs (α_u, ϕ_u) with $\epsilon_u \to 0$,

$$\|I_u - I_0\|_{L^\infty} < \epsilon_u, \qquad \|DI_u\|_{L^\infty} < K, \qquad \|D^2 I_u\|_{L^\infty} < K, \tag{11.86}$$

$$\|g_u - g_0\|_{L^\infty} < \epsilon, \qquad \|Dg_u\|_{L^\infty} < K, \tag{11.87}$$

$$\|\alpha_u\|_{L^\infty} < \epsilon_u, \qquad \|d_{\alpha_u}\phi_u\|_{L^\infty} < \epsilon_u, \tag{11.88}$$

and such that defining ψ_u (resp. $\phi_{u,0}$) as ψ (resp. ϕ_0) in Theorem 11.1 we have the estimate

$$\|d\alpha_u\|_{L^2} < \epsilon_u \|\phi_{u,0}\|_{L^2}, \tag{11.89}$$

and the inequality oposite to (11.84) is satisfied

$$\|\phi_{u,0}\|^2_{L^2(Z_{II})} > \frac{1}{e + e^{-1}} \left(\|\phi_{u,0}\|^2_{L^2(Z_I)} + \|\phi_{u,0}\|^2_{L^2(Z_{III})} \right). \tag{11.90}$$

Define

$$\phi_{u,\mathrm{av}}(t,\theta) := \phi_u(t,\theta) - \frac{1}{2\pi} \int \phi_u(t,\nu) d\nu.$$

It follows from (11.88) that $\|\phi_{u,0}\|_{L^\infty} \to 0$ (unless we specify something different, all limits in this proof will implicitly mean as u goes to ∞). Combining this fact with Lemma 11.8 and (11.90) we deduce that for big enough u we also have

$$\|\phi_{u,\mathrm{av}}\|^2_{L^2(Z_{II})} > \frac{1}{e^{1.5} + e^{-1.5}} \left(\|\phi_{u,\mathrm{av}}\|^2_{L^2(Z_I)} + \|\phi_{u,\mathrm{av}}\|^2_{L^2(Z_{III})} \right). \tag{11.91}$$

Lemma 11.8 combined with (11.89) also implies that

$$\|d\alpha_u\|_{L^2} / \|\phi_{u,\mathrm{av}}\|_{L^2} \to 0. \tag{11.92}$$

Let

$$x_u := \frac{1}{2\pi} \int \phi_u(0,\nu) d\nu.$$

Using a real linear transformation $\mathbb{C}^n \to \mathbb{C}^n$ we may assume that $I_u(x_u) = I_0$ and $g_u(x_u) = g_0$. Furthermore, since $\|I_u - I_0\|_{L^\infty} \to 0$, these linear transformations are uniformly bounded and (maybe after increasing K slightly) they preserve the bounds on DI, $D^2 I$ and Dg.

Using a gauge transformation we can assume that α_u in balanced temporal gauge, so that $\alpha_u = a_u d\theta$ for some function $a_u : C \to \mathbf{i}\mathbb{R}$ and the restriction of α_u to $\{0\} \times S^1$ is equal to a constant $\lambda_u \in \mathbf{i}\mathbb{R}$. Then we have $\alpha_u = \lambda_u d\theta + \beta_u d\theta$, where

$$|\lambda_u| < \|\alpha_u\|_{L^\infty} \qquad \text{and} \qquad \|\beta_u\|_{L^\infty} \leq 2\|d\alpha_u\|_{L^\infty}. \tag{11.93}$$

Let us define $\xi_u(t,\theta) := \phi_u(t,\theta) + \mathbf{i}\lambda_u t I_0 \mathcal{X} - x_u$ and

$$\xi_{u,\mathrm{av}}(t,\theta) := \xi_u(t,\theta) - \frac{1}{2\pi}\int \xi_u(t,\nu)d\nu. \tag{11.94}$$

Then we have $\xi_{u,\mathrm{av}} = \phi_{u,\mathrm{av}}$. It follows that $\|d\xi_u\|_{L^2} > 0$, for otherwise we would have $\xi_{u,\mathrm{av}} = \phi_{u,\mathrm{av}} = 0$, in contradiction with (11.91).

Lemma 11.11. *The following holds:* $\|\overline{\partial}_{I_0}\xi_u\|_{L^2}/\|d\xi_u\|_{L^2} \to 0$.

Proof. Applying (11.82) to the pair (α_u, ϕ_u) we compute:

$$\begin{aligned}\overline{\partial}_{I_0}\xi_u = \frac{\partial \xi_u}{\partial t} - I_0\frac{\partial \xi_u}{\partial \theta} &= (I_u(\phi_u) - I_0)\left(\frac{\partial \phi_u}{\partial \theta} - \mathbf{i}\lambda_u\mathcal{X}\right) - \mathbf{i}\beta_u I_u(\phi_u)\mathcal{X} \\ &= (I_u(\xi_u + x_u) - I_u(x_u))\left(\frac{\partial \phi_u}{\partial \theta} - \mathbf{i}\lambda_u\mathcal{X}\right) - \mathbf{i}\beta_u I_u(\xi_u + x_u)\mathcal{X},\end{aligned} \tag{11.95}$$

where in the second equality we have used the fact that I_u is invariant under translations along $\mathcal{X}$. We claim that for some constant K' depending on K

$$\|I_u(\xi_u + x_u) - I_u(x_u)\|_{L^2} < K'\|d\xi_u\|_{L^2}. \tag{11.96}$$

Indeed, using (11.86) we deduce that for any $x \in \mathbb{C}^n$ at distance less than 1 from x_u we have, for some K' depending on K, $\|I_u(x) - I_u(x_u) - DI(x - x_u)\| < K'\|x - x_u\|^2$. Applying this to $x = \xi_u + x_u$ and using the fact that $\|d\xi_u\|_{L^\infty}$ is uniformly bounded (which follows from (11.88) and implies that $\xi_u(C)$ stays uniformly not too far from x_u) we get (11.96). The estimates (11.88) imply that $\|d\phi_u\|_{L^\infty} \to 0$, which on its turn implies

$$\left\|\frac{\partial \phi_u}{\partial \theta}\right\|_{L^\infty} \to 0. \tag{11.97}$$

Also, combining (11.88) with (11.93) we deduce that

$$\lambda_u \to 0. \tag{11.98}$$

Finally, combining (11.92) with $\|\phi_{u,\mathrm{av}}\|_{L^2} = \|\xi_{u,\mathrm{av}}\|_{L^2} < K'\|d\xi_u\|_{L^2}$ we deduce

$$\beta_u/\|d\xi_u\|_{L^2} \to 0. \tag{11.99}$$

Combining (11.95) with (11.96)—(11.99), we obtain the desired limit. $\square$

Now define $\xi'_u := \xi_u/\|d\xi_u\|_{L^2}$. Since $\int \xi'_u(0,\nu)d\nu = 0$ and $\|d\xi'_u\|_{L^2} = 1$, the L^2 norm of ξ'_u is uniformly bounded above, so there is a constant K_0 such that

$$1 \le \|\xi'_u\|_{L^2_1} \le K_0.$$

Since the inclusion $L^2_1 \subset L^2$ is compact, it follows that, up passing to a subsequence, there is some nonzero $\xi \in L^2(C, \mathbb{C}^n)$ such that $\xi'_u \to \xi$ in L^2. We claim that ξ is in fact holomorphic in the interior of C. Indeed, for any test function g supported in the interior of C we have $\langle \xi, \overline{\partial}^* g\rangle_{L^2} = \lim\langle \xi'_u, \overline{\partial}^* g\rangle_{L^2} = \lim\langle \overline{\partial}\xi'_u, g\rangle_{L^2} = 0$. Hence ξ is a weak solution of $\overline{\partial} = 0$, and from standard regularity results we deduce that ξ is smooth and a strong solution: $\overline{\partial}\xi = 0$. Combining this with Gårding's inequality (in the interior of

C): $\|\xi'_u - \xi\|_{L^2_1} \leq K'(\|\overline{\partial}(\xi'_u - \xi)\|_{L^2} + \|\xi'_u - \xi\|_{L^2})$ we deduce that $\|\xi'_u - \xi\|_{L^2_1} \to 0$, so ξ'_u converges to ξ in L^2_1 and hence

$$\|d\xi\|_{L^2} = 1. \tag{11.100}$$

Define $\xi'_{u,\mathrm{av}}$ and ξ_{av} exactly as we defined $\xi_{u,\mathrm{av}}$ in (11.94). Since $\xi'_u \to \xi$ in L^2, it follows that $\xi'_{u,\mathrm{av}} \to \xi_{\mathrm{av}}$ in L^2. On the other hand, since $\xi'_{u,\mathrm{av}}$ is a rescaling of $\xi_{u,\mathrm{av}}$ and we have $\xi_{u,\mathrm{av}} = \phi_{u,\mathrm{av}}$, the formula (11.91) implies, passing to the limit, that

$$\|\xi_{\mathrm{av}}\|^2_{L^2(Z_{II})} \geq \frac{1}{e^{1.5} + e^{-1.5}} \left(\|\xi_{\mathrm{av}}\|^2_{L^2(Z_I)} + \|\xi_{\mathrm{av}}\|^2_{L^2(Z_{III})}\right). \tag{11.101}$$

This contradicts (2) in Lemma 10.10 (applied to $\phi = \xi$) unless ξ is constant, which is impossible by (11.100). This finishes the proof of (1).

11.7.2. *Proof of (2).* We follow the same scheme of proof as before, assuming the existence of sequences I_u, g_u, I_u-holomorphic pairs (α_u, ϕ_u), and $\epsilon_u \to 0$ satisfying (11.86)—(11.88) and

$$\|d\alpha\|_{L^2_1} < K, \qquad \|d\alpha\|_{L^\infty} < K, \tag{11.102}$$

but contradicting (11.85). We will see that this is not possible. Define ξ_u and $\xi_{u,\mathrm{av}}$ as before. We state two lemmae whose proof will be given below so as not to break the argument. The proof of the Lemma 11.12 will be given in Section 11.8 and that of Lemma 11.13 in Section 11.9.

Lemma 11.12. *For big enough u we have $\|\xi_u\|_{L^2} < 2\|\xi_{u,\mathrm{av}}\|_{L^2} = 2\|\phi_{u,\mathrm{av}}\|_{L^2}$.*

Lemma 11.13. *There is some constant K_0 such that, for big enough u, $\|\xi_u\|_{L^2_2} \leq K_0$.*

Let $\|\xi_u\|_{1/2}$ denote the Hölder $C^{0+\frac{1}{2}}$ norm of ξ_u. Since in real dimension 2 the Sobolev space L^2_2 is included in $C^{0+\frac{1}{2}}$, Lemma 11.13 gives a uniform bound (here K_0 may increase from one line to the other, but will always be independent of u)

$$\|\xi_u\|_{1/2} < K_0.$$

Taking this into account and applying Lemma 11.14 to ξ_u we obtain for big enough u

$$\sup_Z |\xi_u| \leq K_0(\|\xi_u\|^{1/4}_{L^2} + \|\xi_u\|_{L^2}) \leq K_0\|\xi_u\|^{1/4}_{L^2}, \tag{11.103}$$

(recall that $\|\xi_u\|_{L^2}$ goes to 0). It is easy to check that $\sup_Z |\xi_{u,\mathrm{av}}| \leq K_0 \sup_Z |\xi_u|$. Combining this observation with the equality $\xi_{u,\mathrm{av}} = \phi_{u,\mathrm{av}}$ and (11.79) in Lemma (11.8) we deduce

$$\sup_Z |\phi_{u,0}| \leq K_0 \sup_Z |\xi_u|. \tag{11.104}$$

On the other hand, combining Lemma 11.12 with (11.78) in Lemma 11.8 we conclude

$$\|\xi_u\|_{L^2} \leq K_0\|\phi_{u,0}\|_{L^2}. \tag{11.105}$$

Putting together (11.103)—(11.105) we deduce $\sup_Z |\phi_{u,0}| \leq K_0\|\phi_{u,0}\|^{1/4}_{L^2}$, so (11.85) must hold, in contradiction with our assumption. This finishes the proof of (2).

Lemma 11.14. *There is a constant $K_0 > 0$ with the following property. Let $f : C \to \mathbb{C}^n$ be a smooth map and let $\|f\|_{1/2}$ denote its Hölder $C^{0+\frac{1}{2}}$ norm. Then*

$$\sup_Z |f| \le K_0(\|f\|_{L^2}^{1/4}\|f\|_{1/2}^{4/5} + \|f\|_{L^2}). \tag{11.106}$$

Proof. Denote for convenience $H := \|f\|_{1/2}$, and let $x \in Z$ be such that $|f(x)| = \sup |f|$. Let

$$\rho = \frac{|f(x)|^2}{4H^2}.$$

Then for any $y \in C$ such that $|x-y| \le \rho$ we have $|f(y)| \ge |f(x)|/2$, because

$$\frac{|f(x)-f(y)|}{\rho^{1/2}} \le \frac{|f(x)-f(y)|}{|x-y|^{1/2}} \le H \implies |f(y)| \ge |f(x)| - \rho^{1/2}H \ge |f(x)| - \frac{|f(x)|}{2}.$$

Now suppose that $\rho \le 1$, so that the ball B of radius ρ centered at x is contained in C. Then $\|f\|_{L^2} \ge \|f\|_{L^2(B)} \ge K_0\rho^2|f(x)|$, which implies, rearranging, that

$$|f(x)| \le K_0\|f\|_{L^2}^{1/5}\|f\|_{1/2}^{4/5}. \tag{11.107}$$

On the other hand, if $\rho > 1$ we estimate $|f(x)| \le K_0\|f\|_{L^2}$. Summing the two inequalities we obtain (11.106). □

11.8. **Proof of Lemma 11.12.** Pick some big u and let $f := \xi_u$ and $f_{\mathrm{av}} := \xi_{u,\mathrm{av}}$. By definition we have

$$|f(t,\theta) - f_{\mathrm{av}}(t,\theta)| = \frac{1}{2\pi}\int f(t,\nu)d\nu = \frac{1}{2\pi}\int_0^t\int \frac{\partial f}{\partial t}(\tau,\nu)d\nu d\tau, \tag{11.108}$$

since $\int f(0,\nu)d\nu = 0$. On the other hand, denoting $\overline{\partial}_{I_0}$ by $\overline{\partial}$ we have

$$\frac{\partial f}{\partial t}(t,\nu) = I_0\frac{\partial f}{\partial \theta}(t,\nu) + \overline{\partial}f(t,\nu).$$

Integrating for $\theta \in S^1$ the first term in the right hand side vanishes, so we obtain

$$\left|\int_0^t\int \frac{\partial f}{\partial t}(\tau,\nu)d\nu d\tau\right| \le K\|\overline{\partial}f\|_{L^2} \tag{11.109}$$

for some constant K. Let K' be the constant in Gårding's inequality

$$\|df\|_{L^2} \le \|f\|_{L^2_1} \le K'(\|\overline{\partial}f\|_{L^2} + \|f\|_{L^2}).$$

Using Lemma 11.11 we know that, if u is big enough, $\|\overline{\partial}f\|_{L^2} < 1/2K'\|df\|_{L^2}$. Rearranging the terms in the inequality, this implies that $\|df\|_{L^2} < 2K'\|f\|_{L^2}$. Using again Lemma 11.11 we conclude that for big enough u we have $\|\overline{\partial}f\|_{L^2} < (2K\,\mathrm{Vol}(C))^{-1}\|f\|_{L^2}$. Taking $L^2(C)$ norms in (11.108) and combining the previous inequality with (11.109) we obtain

$$\|f\|_{L^2} \le \|f_{\mathrm{av}}\|_{L^2} + \frac{1}{2}\|f\|_{L^2}.$$

Hence, $\|f\|_{L^2} \le 2\|f_{\mathrm{av}}\|_{L^2}$, which is what we wanted to prove.

11.9. **Proof of Lemma 11.13.** As before, throughout this proof K_0 will denote a positive number which may increase from line to line but which will always be independent of u. We estimate the L^2_1 norm of $\overline{\partial}_{I_0}\xi_u$ using formula (11.95). First, since DI is uniformly bounded and the diameter of $\xi_u(C)$ tends to 0 (because of (11.88)), we have $\|I_u(\xi_u+x_u)-I_u(x_u))\|_{L^\infty}\to 0$. Hence, for any δ and big enough u we have

$$\begin{aligned}\left\|(I_u(\xi_u+x_u)-I_u(x_u))\frac{\partial\xi_u}{\partial\theta}\right\|_{L^2_1} &\le \|I_u(\xi_u+x_u)-I_u(x_u))\|_{L^\infty}\|\xi_u\|_{L^2_2}\\ &\quad+\|I_u(\xi_u+x_u)-I_u(x_u))\|_{L^2_1}\|\xi_u\|_{L^\infty}\\ &\le \delta\|\xi_u\|_{L^2_2}+\|I_u(\xi_u+x_u)-I_u(x_u))\|_{L^2_1}\|\xi_u\|_{L^\infty}.\end{aligned}$$

(Note by the way that $\partial\xi_u/\partial\theta=\partial\phi_u/\partial\theta$.) Now we can estimate

$$\|I_u(\xi_u+x_u)-I_u(x_u))\|_{L^2_1}\le K_0\|\xi_u\|_{L^2_1}$$

We also know that if u is big enough $\|\xi_u\|_{L^2_1}\le 2\|\xi_u\|_{L^2}$ (this follows from Lemma 11.11 and Gårding's inequality, see for example the argument in Section 11.8). Furthermore, $\|\xi_u\|_{L^2}$ goes to 0. Also, since $\int\xi_u(0,\nu)d\nu=0$ and the diameter of $\xi_u(C)$ goes to 0, $\|\xi_u\|_{L^\infty}<K_0$ holds if u is big enough. Putting these observations together we obtain

$$\left\|(I_u(\xi_u+x_u)-I_u(x_u))\frac{\partial\xi_u}{\partial\theta}\right\|_{L^2_1}\le\delta\|\xi_u\|_{L^2_2}+K_0. \tag{11.110}$$

For the next term in (11.95) the following is easy to prove (if u is big enough), taking into account the previous arguments:

$$\|(I_u(\xi_u+x_u)-I_u(x_u))\mathbf{i}\lambda_u\mathcal{X}\|_{L^2_1}\le K_0\|\xi_u\|_{L^2_1}\le K_0. \tag{11.111}$$

The remaining term satisfies, provided u is big enough,

$$\begin{aligned}\|\beta_uI_u(\xi_u)\mathcal{X}\|_{L^2_1}&\le K_0(\|\beta_u\|_{L^2_1}+\|\beta_u\|_{L^\infty}\|\xi_u\|_{L^2_1})\\ &\le K_0(\|d\alpha_u\|_{L^2_1}+K\|d\alpha_u\|_{L^\infty}\|\xi_u\|_{L^2_1})\\ &\le K_0(\|d\alpha_u\|_{L^2_1}+\|\xi_u\|_{L^2_1})\le K_0,\end{aligned} \tag{11.112}$$

where here we use (11.102) and the inequality $\|\beta_u\|_{L^2_1}\le K_0\|d\alpha_u\|_{L^2_1}$, which is easy to prove. Taking this into account and combining (11.110), (11.111) and (11.112), formula (11.95) implies that $\|\overline{\partial}_{I_0}\xi_u\|_{L^2_1}\le\delta\|\xi_u\|_{L^2_2}+K_0$. Finally, using the standard inequality

$$\|\xi_u\|_{L^2_2}\le K'(\|\overline{\partial}_{I_0}\xi_u\|_{L^2_1}+\|\xi_u\|_{L^2_1})$$

and taking δ smaller than $1/(2K')$ we obtain the desired bound.

12. Limits of approximate gradient lines

Recall that $H=-\mathbf{i}\mu$. We denote for convenience $V:=I\mathcal{X}$, so that V is the negative gradient of H.

Theorem 12.1. *Let $\sigma > 0$ be a real number. Suppose that $\{\psi_u : T_u \to X, l_u, G_u\}$ is a sequence of triples in which each ψ_u is a smooth map with domain a finite closed interval $T_u \subset \mathbb{R}$, each l_u is a nonzero real number and each $G_u > 0$ is a real number. Suppose that for each u and $t \in T_u$ we have*

$$|\psi_u'(t) - l_u V(\psi_u(t))| \leq G_u e^{-\sigma d(t,\partial T_u)}. \tag{12.113}$$

Suppose also that $G_u \to 0$ and that $l_u \to 0$. Passing to a subsequence, we can assume that $l_u|T_u|$ converges somewhere in $\mathbb{R} \cup \{\pm\infty\}$ (here $|T_u|$ denotes the length of T_u). Then we have the following.

1. *If $\lim l_u|T_u| = 0$ then $\lim \operatorname{diam} \psi_u(T_u) = 0$.*
2. *If $\lim l_u|T_u| \neq 0$, define for big enough u and for every $t \in S_u$ the rescaled objects $S_u := l_u T_u$ and $f_u(t) := \psi_u(t/l_u)$. There is a subsequence of $\{f_u, S_u\}$ which converges to a chain of gradient segments $\mathcal{T}$ in X.*

Proof. The case $l_u|T_u| \to 0$ is obvious, so we consider the case $\lim l_u|T_u| > 0$. We begin modifying slightly the definition of S_u. Taking u big enough we can assume that $l_u < 1$. Suppose that $T_u = [a, b]$ and let $T_u' := [a, a - \ln l_u/\sigma]$ and $T_u'' := [b + \ln l_u/\sigma, b]$. Using (12.113) we can bound

$$\begin{aligned} \operatorname{diam}(\psi_u(T_u')) &\leq \int_{T_u'} l_u|V|dt + G_u \int_0^{-\ln l_u/\sigma} e^{-\sigma x} dx \\ &\leq l_u(\sup|V|)|\ln l_u|/\sigma + G_u(1 - l_u)/\sigma \to 0, \end{aligned}$$

and similarly $\operatorname{diam}(\psi_u(T_u'')) \to 0$. Consequently, if we define

$$S_u := [(a - \ln l_u/\sigma)/l_u, (b + \ln l_u/\sigma)/l_u] \qquad \text{and} \qquad f_u(t) := \psi(t/l_u), \tag{12.114}$$

then the statement of the theorem is equivalent to saying that the sequence (f_u, S_u) defined by (12.114) has a subsequence converging to a chain of gradient segments $\mathcal{T}$. Furthermore, equation (12.113) implies that for any $t \in S_u$ we have

$$|f_u'(t) - V(f_u(t))| \leq G_u e^{-\sigma d(t,\partial S_u)/l_u} \leq G_u e^{-\sigma d(t,\partial S_u)}. \tag{12.115}$$

Lemma 12.2. *For any connected component $F_0 \subset F$ of the fixed point set and any small enough $\delta > 0$ there exist numbers $0 < d < \delta$, η and $c > 0$ with the following property: let $f : [a, b] \to X$ be a map satisfying $|f'(t) - V(f(t))| \leq \eta$ for every t and, for some $a < t < t' < b$,*

$$f(t) \in F_0^d \qquad \text{and} \qquad f(t') \notin F_0^{2\delta}$$

(recall that F_0^d and $F_0^{2\delta}$ denote the d and 2δ-neighbourhoods of F_0 respectively). Then, for every $t'' > t' + \delta$ we have

$$H(f(t'')) \leq H(F_0) - c.$$

In particular, if d is small enough the for every $t'' > t' + \delta$ we have $f(t'') \notin F_0^d$ (i.e., f never comes back to F_0^d after time $t' + \delta$).

The proof of Lemma 12.2 will be given in Section 12.1. Now we continue with the proof of Theorem 12.1. Passing to a subsequence, we can assume that there is some $K > 0$ such that for any connected component $F' \subset F$ either $\lim_{u\to\infty} d(\psi_u(S_u), F') = 0$ or $d(\psi_u(S_u), F') \geq K$ for every u. Let $F_1, \dots, F_l \subset F$ be the connected components which fall in the first case. We claim that the values of H in each of the components $F_1, \dots, F_l$ are all different.

To prove this, suppose on the contrary that for some i and j we have $H = H(F_i) = H(F_j)$. Take some small δ, and let $0 < d < \delta$, η and c be the numbers obtained by taking $F_0 := F_i$ in Lemma 12.2. Let also $\alpha > 0$ be so small so that $H(F_j^{\alpha}) \subset [H-c/2, H-c+2]$. Then, if G_u is smaller than η, and for some t we have $f_u(t) \in F_i^d$, then for every $t' \geq t$ we have $f_u(t) \notin F_j^{\alpha}$. Interchanging the roles of F_i and F_j we prove in the same way that there are some d', α' such that if $f_u(t) \in F_j^{d'}$ and $t' \geq t$ then $f_u(t') \notin F_i^{\alpha'}$. So we cannot have simultaneously $\lim_{u\to\infty} d(\psi_u(S_u), F_i) = 0$ and $\lim_{u\to\infty} d(\psi_u(S_u), F_j) = 0$.

Hence we can suppose that $H(F_1) > H(F_2) > \cdots > H(F_l)$. Now we pick some very small δ. Suppose that u is so big that we can apply Lemma 12.2 to each of the components $F_1, \dots, F_l$. Let $d_1, \dots, d_l$ be the numbers given by Lemma 12.2 for our choice of δ. Then we define, for every $1 \leq j \leq l$,

$$E_{u,j}^{\delta} := [\inf\{t \in S_u \mid f(t) \in F_j^{d_j}\}, \inf\{t \in S_u \mid f(t) \notin F_k^{2\delta}\} + \delta].$$

Let E_u^{δ} be the union $E_{u,1}^{\delta} \cup \cdots \cup E_{u_l}^{\delta}$ and let T_u^{δ} denote the closure of the complementary $S_u \setminus E_u^{\delta}$.

Passing to a subsequence we can assume that for all δ and u all sets T_u^{δ} have the same number of connected components: $T_u^{\delta} = T_{u,1}^{\delta} \cup \cdots \cup T_{u,p}^{\delta}$, where p lies between $l-1$ and $l+1$, and that the inequality $T_{u,1}^{\delta} \leq E_{u,1}^{\delta}$ holds either for all u or for none of them; if the inequality holds and furthermore

$$\limsup_{\delta\to 0} \limsup_{u\to\infty} \operatorname{diam}(f_u(T_{u,0}^{\delta})) = 0,$$

then we remove $T_{u,1}^{\delta}$ from T_u^{δ} and attach it to $E_{u,1}^{\delta}$. Similarly, we remove $T_{u,p}^{\delta}$ from T_u^{δ} if the diameter of its images converges to 0. After this operations we end up with a new set $T_u^{\delta} = T_{u,1}^{\delta} \cup \cdots \cup T_{u,k}^{\delta}$ for each u and δ.

It follows from the definition of $E_{u,j}^{\delta}$ and from Lemma 12.2, is that for any $t \in T_u^{\delta}$ the point $f(t)$ lies away from the set $X' = F_1^{d_1} \cup \cdots \cup F_l^{d_l}$. Since $\sup_{X\setminus X'} |V| > 0$, it follows that if G_u is small enough then the size of $f'(t)$ is comparable to that of V. As a consequence, for fixed δ the intervals $T_{u,j}^{\delta}$ have bounded length. From this it is rather straightforward to prove, passing to a subsequence, that (6.21) holds and that there is a limiting gradient segment $(x_j^{\delta}, T_j^{\delta})$ to which the images of $T_{u,j}^{\delta}$ converge.

Hence we only need to prove that the images of the intervals $E_{u,j}^{\delta}$ accumulate near the fixed point component F_j and that their diameters tend to 0. The former is almost obvious from the definition, whereas the latter is a bit more subtle and follows from next lemma, using the fact that $f_u(E_{u,j}^{\delta}) \subset F_j^{2\delta}$. □

Lemma 12.3. *There is a constant $K > 0$ with the following property. For any small enough $\delta > 0$, any big enough u (depending on δ), and any interval $E \subset S_u$ such that $f_u(E) \subset F^\delta$, we have* $\operatorname{diam}(f_u(E)) \leq \delta K$.

Proof. If δ is small enough we can assume that $f_u(E)$ lies in the δ-neighbourhood F_0^δ of a unique connected component $F_0 \subset F$. Let N be the normal bundle of F_0. Using the exponential map with respect to the S^1-equivariant metric g, we can identify F_0^δ with a neighbourhood N^δ of the zero section of N. Then we can pullback the vector field V to a vector field on N^δ, which we denote by the same symbol V. We also think of $f_u|_E$ as taking values in N^δ. Consider on N the restriction of the metric g. This gives an equivariant Euclidean metric on N.

The bundle N carries a linear action of S^1, and we can split $N = N_1 \oplus \cdots \oplus N_k$ in such a way that S^1 acts on N_j with weight $w_j \neq 0$. Let $V_0 \in \Gamma(TN^{\mathrm{vert}})$ be the vertican tangent field whose value at a vector $x = (x_1, \ldots, x_k) \in N$ is $V_0(x) := (w_1 x_1, \ldots, w_k x_k)$ (here we are identifying $TN^{\mathrm{vert}} \simeq N$). It is well known that V_0 approximates at first order V near F_0 (which we view as the zero section of N). More precisely, there is a constant K such that for any $x \in N^\delta$ we have

$$|V_0(x) - V(x)| \leq K|x|^2. \tag{12.116}$$

Let $d : E \to \mathbb{R}$ be the function defined as $d(t) := |f_u(t)|^2$. We want to prove that $d(t)$ decays exponentially as t goes away from the extremes of E. We will follow the same idea as in the proof of formula (11.60) in Theorem 11.1; however, in this situation the analysis will be much simpler.

Shifting E (and modifying accordingly f_u) and removing if necessary a small interval of length < 2 at the end of E, we can assume that $E = [-L, L]$, where L is a natural number (it is clear that the truncation does not affect the estimate). Define for every natural number $-L \leq n < L$ the energy $d_n := \int_n^{n+1} |f_u(t)| dt$. Define also $g_n := \int_n^{n+1} |f_u'(t) - V(f_u(t))| dt$.

Lemma 12.4. *Let $\gamma := 1/(e^{1/2} + e^{-1/2})$. If $\delta > 0$ is small enough, then there is some $\epsilon > 0$ such that, for every $-L < n < L - 1$ satisfying $g_{n-1} + g_n + g_{n+1} \leq \epsilon d_n$, we have*

$$d_n \leq \gamma(d_{n-1} + d_{n+1}).$$

The proof of Lemma 12.4 will be given in Section 12.2. We now finish the proof of Lemma 12.3. Inequality (12.115) implies that $g_n \leq G_u e^{-\sigma(L-n)}$. Arguing exactly as in the proof of Lemma 10.9, we deduce that for some $\beta > 0$ independent of f_u there is a bound $d_n \leq (\epsilon^{-1} K G_u + d_{-L} + d_L) e^{-\beta(L-|n|)}$. Then we have (using the fact that $|V(x)| \leq K|x|$)

$$\begin{aligned}
\operatorname{diam}(f_u(E)) &\leq K \int_E |f_u'| \leq K \int_E (|V(f_u)| + |f_u' - V(f_u)|) \\
&\leq K \int |f_u| + K \int_E G_u e^{-\sigma(t, \partial S_u)} dt \leq K \sum_{n=-L}^{L} d_n + K G_u \\
&\leq K(G_u + d_{-L} + d_L) \leq K(G_u + \delta).
\end{aligned}$$

Taking u big enough so that $G_u < \delta$, the result follows. □

12.1. **Proof of Lemma 12.2.** We can assume without loss of generality that $\eta \leq 1$. Define $M := \sup |V| + 1$. Then for every t we have

$$|f'(t)| \leq M. \tag{12.117}$$

Let $h := H(f)$ and let $h_0 := H(F_0)$. We have $h' = -\langle V(f), f' \rangle$, so that

$$|h'(t) + V(f(t))|^2 \leq \eta |V(f(t)|. \tag{12.118}$$

This implies the following: for every $\delta > 0$ there exists some $K > 0$ such that, if $|\eta| \leq K/2$ and $f(t) \in X^{\delta}$ then $h'(t) \leq K/2$.

Lemma 12.5. *Given any small $\epsilon > 0$, if η is small enough (depending on ϵ) we have, for any t,*

(1) *if $f(t) \in X^{\epsilon M + \delta}$ and $h(t) \leq h_0 + \epsilon K/4$ then, for every $t' \geq t + \epsilon$, we have $h(t') \leq h_0 - \epsilon K/4$.*
(2) *if $h(t) \leq h_0 - \epsilon K/4$, then for any $t' \geq t$ we have $h(t') \leq h_0 - \epsilon K/4$;*
(3) *if $h(t) \leq h_0 + \epsilon K/4$, then for any $t' \geq t$ we have $h(t') \leq h_0 + \epsilon K/4$.*

Proof. Suppose that η is so small that whenever $f(t) \in X^{\delta}$ we have $h'(t) \leq K/2$. It follows from (12.117) that if $f(t) \in X^{\epsilon M + \delta}$, then for every $\tau \in [t, t+\epsilon]$ we have $f(\tau) \in X^{\delta}$, so $h'(\tau) \leq K/2$. Integrating we obtain (1). We now prove (2). If ϵ is small enough then V does not vanish in $\{H = h_0 - \epsilon K/4\}$. Taking η smaller than one half of the supremum of $|V|$ on the level set $\{H = h_0 - \epsilon K/4\}$ we deduce from (12.118) that for any τ such that $f(\tau) = h_0 - \epsilon K/4$ we have $h'(\tau) < 0$. This clearly implies (2). The same argument proves (3). □

Take any ϵ satisfying the hypothesis of the previous lemma and also $\epsilon \leq M^{-1}\delta$. Take $d < \delta$ small enough so that we have an inclusion

$$F_0^d \subset \{|H - h_0| \leq \epsilon K/4\}.$$

Define also $c := \epsilon K/4$. We claim that this choice of d and c satisfies the requirements of the lemma. Indeed, suppose that for some $t < t'$ we have $f(t) \in F_0^d$ and $f(t') \notin F_0^{2\delta}$. Then $h(t) \leq h_0 + \epsilon K/4$, so by (3) in Lemma 12.5 we also have $h(t') \leq h_0 + \epsilon K/4$. If δ is small enough so that the 2δ-neighbourhoods of each connected component of F are all disjoint, then we can assume that $f(t') \in X^{2\delta} \subset X^{\epsilon M + \delta}$. Combining this fact with the bound on $h(t')$ and (1) in Lemma 12.5 we deduce that $h(t' + \epsilon) \leq h_0 - K/4$. Finally, (2) in Lemma 12.5 implies that for every $t'' \geq t' + \epsilon$ we have $h(t'') \leq h_0 - K/4$, which is what we wanted to prove (recall that $M \geq 1$, so $\epsilon \leq M^{-1}\delta \leq \delta$).

12.2. **Proof of Lemma 12.4.** Suppose that $f_u|_E$ takes values on the restriction N' of N to the connected component $F' \subset F$ of the fixed point set. Let $N' = N_1 \oplus \cdots \oplus N_k$ be the decomposition in weights of the S^1 action, and denote the corresponding weights by $w_1, \dots, w_k$. Let also A denote the endomorphism of N acting on the subbundle N_j as multiplication by w_j.

Define, for every $t \in [-1,2]$, $\phi(t) := f_u(n+t)$, $G(t) := \phi'(t) - V(\phi(t))$ and $H(t) := V(\phi(t)) - V_0(\phi(t))$. Let $(x_1,\dots,x_k)$ be the coordinates of $\phi(0)$. The relation between ϕ and the integral curves of the linear vector field V_0 is given by Duhamel's formula:

$$\begin{aligned}\phi(t) &= e^{tA}\phi(0) + \int_0^t e^{(s-t)A}(G(s)+H(s))ds \\ &= (e^{w_1 t}x_1,\dots,e^{w_k t}x_k) + \int_0^t e^{(s-t)A}(G(s)+H(s))ds.\end{aligned} \tag{12.119}$$

Let $\phi_0(t) := (e^{w_1 t}x_1,\dots,e^{w_k t}x_k)$ and let $\chi(t) := \int_0^t e^{(s-t)A}(G(s)+H(s))ds$. Define also $t_j := \int_{j-n}^{j-n+1} |\phi_0(t)|dt$ for every $j \in \{n-1,n,n+1\}$. One checks easily (see for example the proof of Lemma 10.11) that

$$t_n \le \frac{1}{e+e^{-1}}(t_{n-1}+t_{n+1}). \tag{12.120}$$

By assumption we have

$$\|G(t)\|_{L^1(I)} \le (g_{n-1}+g_n+g_{n+1}) \le \epsilon d_n.$$

On the other hand, using (12.116) and $|\phi| < K\delta$ we obtain

$$\|H(t)\|_{L^1(I)} \le K\left(\int_I |\phi|^2\right) \le K\delta\|\phi\|_{L^1(I)}.$$

Since $e^{(s-t)A}$ is bounded for $s,t \in [-1,2]$, we can bound

$$\|\chi\|_{L^1(I)} \le K(\|G(t)\|_{L^1(I)} + \|H(t)\|_{L^1(I)}) \le K(\epsilon d_n + \delta\|\phi\|_{L^1(I)}).$$

Taking this estimate for I equal to $[n-1,n]$, $[n,n+1]$ or $[n+1,n+2]$ it follows from (12.119) that, for any $j \in \{n_1,n,n+1\}$, the inequality $|d_j - t_j| \le K\delta(\epsilon d_n + \delta d_j)$ holds. Combining this estimate with (12.120), and taking into account that $\gamma > 1/(e+e^{-1})$, we deduce that if δ and ϵ are small enough then we must have

$$d_n \le \gamma(d_{n-1}+d_{n+1}),$$

which is what we wanted to prove.

References

[CT] J. Chen, G. Tian, Compactification of moduli space of harmonic mappings, *Comment. Math. Helv.* **74** (1999) 201–237.

[D] S. Donaldson, *Floer homology groups in Yang-Mills theory*, Cambridge Tracts in Mathematics **147**, Cambridge University Press, Cambridge, 2002.

[DK] S. Donaldson, P. Kronheimer, *The Geometry of Four-Manifolds*, Oxford Math. Monographs, Oxford University Press, 1990.

[F] A. Floer, Symplectic fixed points and holomorphic spheres, *Comm. Math. Phys.* **120** (1989), no. 4, 575–611.

[Fr] D. Freed, Classical Chern–Simons theory, *Adv. Math.* **113** (1995) 237—303.

[FO] K.Fukaya, K. Ono, Arnold conjecture and Gromov-Witten invariant, *Topology* **38** (1999) 933–1048.

[GKM] M. Goresky, R. Kottwitz, R. MacPherson, Equivariant cohomology, Koszul duality, and the localization theorem, *Invent. Math.* **131** (1998), no. 1, 25–83.

[GLS] V. Guillemin, E. Lerman, S. Sternberg, Symplectic fibrations and multiplicity diagrams, Cambridge University Press, Cambridge, 1996.

[GS] V.W. Guillemin, S. Sternberg, *Supersymmetry and Equivariant de Rham Theory*, Mathematics Past and Present, Springer (1999).

[IS] S. Ivashkovich, V. Shevchishin, Gromov compactness theorem for J-complex curves with boundary, *Internat. Math. Res. Notices* 2000, **22**, 1167–1206.

[Ko] M. Kontsevich, Enumeration of rational curves via torus actions, *The moduli space of curves* (Texel Island, 1994), 335–368, Progr. Math., **129**, Birkhuser Boston, Boston, MA, 1995.

[LT] J. Li, G. Tian, Virtual moduli cycles and Gromov-Witten invariants of general symplectic manifolds, *Topics in symplectic* 4*-manifolds* (Irvine, CA, 1996), 47–83, First Int. Press Lect. Ser., I, Internat. Press, Cambridge, MA, 1998.

[McDS] McDuff, Salamon, J-holomorphic curves and quantum cohomology, University Lecture Series **6**, AMS, Providence, RI.

[M] I. Mundet i Riera, Hamiltonian Gromov–Witten invariants, *Topology* **42** (2003) 525–553.

[MT] I. Mundet i Riera, G. Tian, Virtual moduli cycles and Hamiltonian Gromov–Witten invariants, in preparation.

[PSS] S. Piunikhin, D. Salamon, M. Schwarz, Symplectic Floer-Donaldson theory and quantum cohomology, in *Contact and symplectic geometry* (Cambridge, 1994), 171–200, Publ. Newton Inst. **8**, Cambridge Univ. Press, Cambridge, 1996.

[R] Y. Ruan, Virtual neighborhoods and pseudo-holomorphic curves, Proceedings of 6th Gkova Geometry-Topology Conference, *Turkish J. Math.* **23** (1999) 161–231.

[RT1] Y. Ruan, G. Tian, A mathematical theory of quantum cohomology, *J. Diff. Geom.* **42** (1995), no. 2, 259–367.

[RT2] Y. Ruan, G. Tian, Higher genus symplectic invariants and sigma models coupled with gravity, *Invent. Math.* **130** (1997) 455–516.

[S] B. Siebert, Gromov-Witten invariants for general symplectic manifolds, preprint `http://arXiv.org/abs/dg-ga/9608005`.

[T] G. Tian, Constructing symplectic invariants. *Quantum cohomology* (Cetraro, 1997), 269–311, Lecture Notes in Math **1776**, Springer, Berlin, 2002.

Departament d'Àlgebra i Geometria, Facultat de Matemàtiques, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain

E-mail address: `mundet@mat.ub.es`

Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton NJ 08544-1000 USA

E-mail address: `tian@math.princeton.edu`